DISCONTINUOUS GALERKIN METHODS AND CASCADING MULTIGRID METHODS FOR INTEGRO-DIFFERENTIAL EQUATIONS

JINGTANG MA

Discontinuous Galerkin Methods and Cascading Multigrid Methods for Integro-Differential Equations

by

© [illegible]

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Department of Mathematics and Statistics

Memorial University of Newfoundland

January 2005 Submitted

St. John's Newfoundland Canada

Discontinuous Galerkin Methods and Cascading Multigrid Methods for Integro-Differential Equations

by

©Jingtang Ma

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Department of Mathematics and Statistics
Memorial University of Newfoundland

January 2004 Submitted

St. John's Newfoundland Canada

Abstract

In this thesis, we focus on the discontinuous Galerkin (DG) methods for the functional integro-differential equations and on the cascading multigrid (CMG) methods for the parabolic PDEs, Volterra integro-differential equations (VIDEs) and Fredholm equations.

We give both a priori and a posteriori error estimates of the DG method for linear, semilinear and nonstandard VIDEs. Furthermore the superconvergence of the mesh-dependent Galerkin method for VIDEs is also considered. The fully discretized DG method for VIDEs is also analyzed. Numerical examples are provided to compare the DG method with the continuous Galerkin (CG) method and the continuous collocation (CC) method. We study the primary discontinuities of several classes of VIDEs with time dependent delays, which include the functional VIDEs of Hale's type, delay VIDEs with weakly singular kernels and delay VIDEs of neutral type (with weakly singular kernels). According to the regularity information established, we construct an adaptive DG method for functional VIDEs of Hale's type.

Two new cascading multilevel algorithms are analyzed to the semi-linear parabolic PDEs and extended to the partial Volterra integro-differential equations (PVIDEs) and the parabolic PDEs with delays. More distinctly the cascading multigrid method could very well solve the Fredholm equations without dealing with the full stiffness matrix directly. Therefore we can save much more computing time. Most im-

portantly, we contribute to the *multigrid arts* by developing an abstract cascading multigrid method in Besov spaces and a discontinuous Galerkin cascading multigrid method. We extend these methods to evolutionary equations and PVIDEs. Finally, we discuss briefly the future works on (partial) VIDEs with blow-up solutions and artificial boundary methods for PVIDEs on unbounded domains.

Acknowledgements

I would like to acknowledge that my supervisor, Professor Hermann Brunner, brought me into the area of numerical analysis with his interesting courses, many good ideas and references which the thesis is based on. He has so broad a knowledge that he could answer my all kinds of mathematical questions. He was able to direct me skillfully to develop a rough idea into a nice research paper. His excellent mathematical investigation strongly impressed me and indicated me how to think. I very appreciate his financial support for my attending to the very important international workshops. In a word, I owe all my success to him.

Many thanks go to Dr. Andy Foster and Dr. Xiaoqiang Zhao as members of my supervisor committee. I am also grateful to Dr. Xingfu Zou and Dr. Lin Wang for their encouragement.

I would like to thank the Department of Mathematics and Statistics, and Professors Herb Gaskill and Bruce Watson, the past and present department heads, for providing me with the necessary facilities for research and a teaching assistant fellowship. I am also grateful to the School of Graduate Studies, Memorial University of Newfoundland for supporting me with a Graduate Fellowship.

I would express my thanks to Professors Houde Han and Lancheng Wu for discussions on the artificial boundary method during their visiting the department. My thanks are also extended to Professor Zhongci Shi for his bringing the papers (Shi and Xu (1998, 1999)) to my attention.

I am especially grateful to my wife, Ms. Cuiye Peng, for her understanding and love. Finally, I would like to express appreciation to my parents for their care.

Contents

List of Tables

List of Figures

Chapter 1

The discontinuous Galerkin method for ODEs: an introduction

The discontinuous Galerkin (DG) method can be traced back to 1973 when Reed and Hill used it to solve the neutron transport problem. The DG method was first analyzed in 1974 by Lesaint and Raviart in the application to ODEs. Since DG methods assume discontinuous approximate solutions, they can be considered as generalizations of finite volume methods. What makes DG methods popular is that they are able to capture the physically relevant discontinuities of the exact solutions without producing spurious oscillations near them. A more detailed history of DG methods for ODEs and PDEs will be presented at the end of this chapter, in Section 1.4.

1.1 The DG method for ODEs

1.1.1 Basic description of the DG method for ODEs

Consider the initial value problem

$$y'(t) = f(t, y(t)), \ t \in I := [0, T], \quad y(0) = y_0, \tag{1.1.1.1}$$

and assume that the (Lipschitz continuous) function $f : I \times \Omega \subset \mathbb{R} \to \mathbb{R}$ is such that (1.1.1.1) possesses a unique solution $y \in C^1(I)$ for all $y_0 \in \Omega$. Let

$$I_h := \{t_n : \ 0 = t_0 < t_1 < \cdots < t_M < \cdots < t_N = T\}$$

be a given mesh on I, and set $I_n := (t_{n-1}, t_n]$, $\overline{I}_n := [t_{n-1}, t_n]$, $h_n := t_n - t_{n-1}$ ($n = 1, \ldots, N-1$); $h := \max\{h_n : \ 1 \leq n \leq N-1\}$ will be called the *diameter* of the mesh I_h. Note that we have, in rigorous notation,

$$t_n := t_n^{(N)}, \ I_n := I_n^{(N)}, \ h_n := h_n^{(N)} \ (n = 1, \ldots, N-1), \ h := h^{(N)}.$$

We will usually suppress this dependence on N, the number of subintervals corresponding to a given mesh I_h, except occasionally in the convergence analysis where $N \to \infty$ with $Nh^{(N)}$ uniformly bounded. At the mesh points the left- and right-sided limits of piecewise continuous functions $\varphi : \ I \to \mathbb{R}$ will be important. They are defined as follows:

$$\varphi_n^+ := \lim_{s \to 0, \ s > 0} \varphi(t_n + s), \ 0 \leq n \leq N-1; \ \varphi_n^- := \lim_{s \to 0, \ s > 0} \varphi(t_n - s), \ 1 \leq n \leq N.$$

The jump across the mesh points is given by $[\varphi]_n := \varphi_n^+ - \varphi_n^-$.

In the DG method, we are looking for an approximate solution of (1.1.1.1) in the finite space

$$\mathcal{V}_N^{(m)} := \{\varphi \in L^2(I) : \ \varphi|_{I_n} \in \mathcal{P}^{(m)}(I_n), \ 1 \leq n \leq N\}, \tag{1.1.1.2}$$

where $\mathcal{P}^{(m)}(I_n)$ denotes the space of all (real) polynomials of degree not exceeding m. We define the DG method for (1.1.1.1) as: Find $Y \in \mathcal{V}_N^{(m)}$ such that

$$B_{DG}(Y, X) = Y_0^- X_0^+, \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{1.1.1.3}$$

where $Y_0^- = y_0$ and

$$B_{DG}(Y,X) := \sum_{n=1}^{M} \int_{I_n} (Y'(t) - f(t,Y))X\,dt + \sum_{n=2}^{M} [Y]_{n-1}X_{n-1}^+ + Y_0^+X_0^+, \quad (1.1.1.4)$$

Note that the exact solution y of (1.1.1.1) satisfies

$$B_{DG}(y,X) = y_0 X_0^+, \quad \forall X \in \mathcal{V}_N^{(m)}.$$

Hence the Galerkin orthogonality property,

$$B_{DG}(y - Y, X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}, \quad (1.1.1.5)$$

holds true. We remark also that the DG method (1.1.1.3) can be interpreted as a time-stepping scheme: For $n = 1, \cdots, N$, find $Y|_{I_n} \in \mathcal{P}^{(m)}(I_n)$, such that

$$\int_{I_n} (Y' - f(t,Y))X\,dt + Y_{n-1}^+X_{n-1}^+ = Y_{n-1}^-X_{n-1}^+, \quad \forall X \in \mathcal{P}^{(m)}(I_n). \quad (1.1.1.6)$$

Here we set $Y_0^- := y_0$.

In order to make the readers capture the basic idea of DG methods for ODEs easily, we shall first introduce the DG method with piecewise constant approximation (i.e., $m = 0$ in (1.1.1.2)); it will be denoted by DG(0). If piecewise linear polynomial approximation ($m = 1$) is used, we write DG(1). For more general DG schemes with high-order polynomial approximation, we refer to Section 1.2. Here we also remark that an analogous analysis holds for systems of ODEs if the products are replaced by the corresponding inner products in $\mathbb{R}^d$ (d denotes the dimension of the systems) [see for example Section 1.2]. Throughout this thesis we define the norm $||\cdot||_J := \sup_{t\in J} |\cdot|$, where J is some compact interval.

1.1.2 A priori error estimate for ODEs

In this section we study the a priori error analysis of DG methods for ODEs.

Theorem 1.1.2.1. *Assume that f has continuous partial derivatives and there is a constant $L > 0$ such that*

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2|, \tag{1.1.2.1}$$

for all $t \in I$ and $y_1,\ y_2 \in \Omega$. Then there exists a constant C, independent of h_n, such that for $0 \leq M \leq N$, the error of DG(m) for (1.1.1.1) satisfies

$$||e||_{[0,t_M]} \leq C \max_{n \leq M} h_n^{m+1} ||y^{(m+1)}||_{\bar{I}_n}, \tag{1.1.2.2}$$

with $C := C(L, t_M)$ and $m = 0,\ 1$.

Proof. If $V \in \mathcal{V}_N^{(m)}$ is determined by $V_0^- = y_0$ and, for $1 \leq n \leq M \leq N$, by

$$\int_{I_n} V' X dt - \int_{I_n} f(t, y(t)) X dt + V_{n-1}^+ X_{n-1}^+ = V_{n-1}^- X_{n-1}^+, \tag{1.1.2.3}$$

for all $X \in \mathcal{V}_N^{(m)}$ $(m = 0,\ 1)$, then

$$||y - V||_{\bar{I}_n} \leq C h_n^{m+1} ||y^{(m+1)}||_{\bar{I}_n}. \tag{1.1.2.4}$$

(See Estep [43] for the proof of (1.1.2.4)). We now compare Y to V. Setting $\mu := y - V$ and $\phi := Y - V \in \mathcal{V}_N^{(m)}$, we have $e = \mu - \phi$. We subtract (1.1.2.3) from (1.1.1.6) and obtain

$$\int_{I_n} \phi' X dt - \int_{I_n} (f(t, Y) - f(t, y)) X(t) dt + \phi_{n-1}^+ X_{n-1}^+ = \phi_{n-1}^- X_{n-1}^+, \tag{1.1.2.5}$$

for all $X \in \mathcal{V}_N^{(m)}$. Choosing $X = \phi$ leads to

$$\frac{1}{2} \left|\phi_{n-1}^+\right|^2 + \frac{1}{2} \left|\phi_n^-\right|^2 - \int_{I_n} (f(t, Y) - f(t, y)) \phi(t) dt = \phi_{n-1}^- \phi_{n-1}^+. \tag{1.1.2.6}$$

Subsequently, we arrive at

$$\frac{1}{2} \left|\phi_n^-\right|^2 \leq \frac{1}{2} \left|\phi_{n-1}^-\right|^2 + \int_{I_n} \left|(f(t, Y) - f(t, y)) \phi(t)\right| dt. \tag{1.1.2.7}$$

We now substitute $e = \mu - \phi$ into (1.1.2.7) to find that

$$|\phi_n^-|^2 \leq |\phi_{n-1}^-|^2 + L \int_{I_n} |\mu|^2 dt + 3L \int_{I_n} |\phi|^2 dt. \tag{1.1.2.8}$$

Next we choose $X = (t - t_{n-1})\phi'$ in (1.1.2.5) to obtain

$$h_n^2 ||\phi'||_{I_n}^2 \leq 8L^2 h_n \int_{I_n} |\mu|^2 dt + 8L^2 h_n \int_{I_n} |\phi|^2 dt.$$

Since

$$\int_{I_n} |\phi|^2 \, dt \leq 2h_n \, |\phi_n^-|^2 + \frac{2}{3} h_n^3 ||\phi'||_{I_n}^2,$$

we find that

$$\int_{I_n} |\phi|^2 \, dt \leq 4h_n \, |\phi_n^-|^2 + \int_{I_n} \mu^2 dt, \tag{1.1.2.9}$$

provided that $\frac{16}{3} L^2 h_n^2 \leq \frac{1}{2}$.

By combining (1.1.2.8) and (1.1.2.9), we see that

$$|\phi_n^-|^2 \leq \frac{1}{1 - 12Lh_n} |\phi_{n-1}^-|^2 + \frac{4L}{1 - 12Lh_n} \int_{I_n} |\mu|^2 dt. \tag{1.1.2.10}$$

Now we iterate (1.1.2.10), assuming $12Lh_n \leq \frac{1}{2}$. This yields

$$|\phi_M^-|^2 \leq CLt_M ||\mu||_{[0,t_M]}^2. \tag{1.1.2.11}$$

Theorem 1.1.2.1 follows from (1.1.2.11) directly when $m = 0$. For $m = 1$, we use equation (1.1.2.6) and the fact that $\phi \in \mathcal{V}_N^{(m)}$ implies that $||\phi||_{I_n}^2 \leq |\phi_{n-1}^+|^2 + |\phi_n^-|^2$, to obtain

$$\frac{1}{4} ||\phi||_{I_n}^2 \leq |\phi_{n-1}^-|^2 + \frac{L}{2} \int_{I_n} \mu^2 dt + \frac{3L}{2} \int_{I_n} \phi^2 dt.$$

An analogous argument establishes Theorem 1.1.2.1 for $m = 1$. We also refer to Estep [43] for the original proof of Theorem 1.1.2.1.

Corollary 1.1.2.1. *Assume $f(t,y) := -a(t)y(t) + g(t)$ in (1.1.1.1) where a, g are continuous on each interval I_n $(1 \le n \le N)$, and set $\mathcal{A} := ||a||_I$. Then there exists a constant C, independent of h, such that the error of DG(m) satisfies*

$$||y - Y||_{[0,t_M]} \le Ch^{m+1}||y^{(m+1)}||_{[0,t_M]},$$

with $C := C(t_M, \mathcal{A})$ and $m = 0,\ 1$.

1.1.3 A posteriori error estimates for ODEs

The a posteriori error analysis is based on representing the error in terms of the solution of a continuous dual problem related to (1.1.1.1), which is used to determine the effects of the accumulation of errors, and in terms of the residual of the computed solution, which measures the propagation of error. After showing the stability of the dual problem, we can estimate the a posteriori error bound of DG(m). The details for the DG(0) method can be found in the book [45].

Theorem 1.1.3.1. *Assume $f(t,y) = -a(t)y(t) + g(t)$ in (1.1.1.1) where a, g are continuous on each interval I_n $(1 \le n \le N)$, and let $m = 0,\ 1$. Then the a posteriori error of DG(m) at the mesh point t_M $(0 \le M \le N)$ satisfies*

$$|y(t_M) - Y_M^-| \le S(t_M)||h_n R(Y)||_{[0,t_M]},$$

where $|R(Y)| := \frac{|Y_{n-1}^+ - Y_{n-1}^-|}{h_n} + |g - aY|$ $(t \in I_n)$ and $\|\cdot\| := \max_{1\le n\le M}\{\|\cdot\|_{I_n}\}$. If $\mathcal{A} := ||a||_I$, then $S(t_M) \le \exp(\mathcal{A}t_M)$ and if , in addition, $a(t) \ge 0$ for all $t \in I$, then $S(t_M) \le 1$.

Proof. The proof for DG(0) is in [45]. The analysis of a posteriori error estimate of DG(1) for (1.1.1.1) is similar to DG(0). We give the details as follows.

Firstly we consider the continuous dual problem for (1.1.1.1): Find $x = x(t)$ such that

$$\begin{cases} -x' + a(t)x = 0, & \text{for } t_M > t \geq 0, \\ x(t_M) = e_M^-, \end{cases} \tag{1.1.3.1}$$

where we denote by $e(t) := y(t) - Y(t)$ the error at the time t and set $e_M^- := y(t_M) - Y_M^-$. Starting from the identity

$$(e_M^-)^2 = (e_M^-)^2 + \sum_{n=1}^{M} \int_{I_n} e \cdot (-x' + a(t)x)dt,$$

we integrate by parts over each subinterval I_n to obtain

$$(e_M^-)^2 = \sum_{n=1}^{M} \int_{I_n} (e' + a(t)e)xdt + \sum_{n=1}^{M-1} [e]_n x_n^+ + (y_0 - Y_0^+)x_0^+. \tag{1.1.3.2}$$

Using $Y_0^- = y_0$, we can simplify (1.1.3.2) to

$$(e_M^-)^2 = \sum_{n=1}^{M} (\int_{I_n} (g - aY - Y')xdt - [Y]_{n-1} x_{n-1}^+).$$

We use the Galerkin orthogonality (1.1.1.5) by choosing $X = \tilde{x}$ to be the L_2 projection into the space $\mathcal{V}_N^{(1)}$ and obtain the error representation formula:

$$(e_M^-)^2 = \sum_{n=1}^{M} (\int_{I_n} (g - aY)(x - \tilde{x})dt - [Y]_{n-1}(x - \tilde{x})_{n-1}^+).$$

From the error estimates of interpolation:

$$\int_{I_n} |x - \tilde{x}|dt \leq h_n \int_{I_n} |x'|dt; \ |x - \tilde{x}| \leq \int_{I_n} |x'|dt,$$

we arrive at

$$\begin{aligned}
(e_M^-)^2 &\leq \sum_{n=1}^{M} \left\{ ||g - aY||_{\bar{I}_n} \int_{I_n} |x - \tilde{x}|dt + \frac{|[Y]_{n-1}|}{h_n} h_n |x - \tilde{x}| \right\} \\
&\leq \sum_{n=1}^{M} \left\{ h_n ||(|g - aY| + \frac{|[Y]_{n-1}|}{h_n})||_{\bar{I}_n} \int_{I_n} |x'|dt \right\} \\
&\leq \max_{1 \leq n \leq M} \left\{ h_n ||(|g - aY| + \frac{|[Y]_{n-1}|}{h_n})||_{\bar{I}_n} \right\} \int_0^{t_M} |x'|dt \\
&\leq S(t_M) \cdot |e_M^-| \cdot ||h_n R(Y)||_{[0,t_M]},
\end{aligned} \tag{1.1.3.3}$$

where $|R(Y)| := |g - aY| + \frac{|[Y]_{n-1}|}{h_n}$. The stability factor $S(t_M)$ is defined by

$$S(t_M) := \frac{\int_0^{t_M} |x'|dt}{|e_M^-|}.$$

For the estimate of $S(t_M)$, we give the following lemma.

Lemma 1.1.3.1. *If $\mathcal{A} := ||a||_I$, then the solution x of (1.1.3.1) satisfies*

$$|x(t)| \leq \exp(\mathcal{A}t_M)|e_M^-|,$$

for all $0 \leq t \leq t_M$ and $S(t_M) \leq \exp(\mathcal{A}t_M)$. If, in addition, $a(t) \geq 0$ for all $t \in I$, then x satisfies

$$|x(t)| \leq |e_M^-|,$$

for all $0 \leq t \leq t_M$ and $S(t_M) \leq 1$.

Proof. The proof can be found in [45].

By combining Lemma 1.1.3.1 and (1.1.3.3) we complete the proof of Theorem 1.1.3.1.

The following theorem gives a measure for the efficiency of the a posteriori estimator in Theorem 1.1.3.1.

Theorem 1.1.3.2. *Assume $f(t,y) = -a(t)y(t) + g(t)$ in (1.1.1.1) and that a, g are continuous on each interval I_n $(1 \leq n \leq N)$, and let $\mathcal{A} := ||a||_I$. Then we have*

$$\begin{aligned} |y(t_M) - Y_M^-| &\leq S(t_M)||h_n R(Y)||_{[0,t_M]} \\ &\leq CS(t_M)(1 + \mathcal{A}t_M e^{C\mathcal{A}t_M})^{1/2} h||y'||_{[0,t_M]}, \end{aligned}$$

for $0 \leq M \leq N$.

Proof. We need to estimate $||hR(Y)||_{[0,t_M]}$ in Theorem 1.1.3.1 and Theorem 1.1.3.1:

$$\begin{aligned} ||h_n R(Y)||_{[0,t_M]} &:= ||\,|Y^+_{n-1} - Y^-_{n-1}| + h_n|g - aY|\,||_{[0,t_M]} \\ &= ||Y^+_{n-1} - y + y - Y^-_{n-1}||_{[0,t_M]} \\ &+ h_n||a(t)y(t) - a(t)Y(t)||_{[0,t_M]} + h||y'||_{[0,t_M]} \\ &\leq (2 + \mathcal{A})||e(t)||_{[0,t_M]} + h||y'||_{[0,t_M]}. \end{aligned} \tag{1.1.3.4}$$

Upon applying Corollary 1.1.2.1, we complete the proof.

1.2 Mesh-dependent Galerkin methods for ODEs

In this section we survey the paper Delfour and Dubeau [40] and discuss the mesh-dependent Galerkin methods (including the discontinuous Galerkin method) for ODEs and the corresponding superconvergence results.

1.2.1 Mesh-dependent Galerkin methods for ODEs

Consider the following system of ODEs,

$$y'(t) = f(t, y(t)), \quad t \in I := [0, T], \ y(0) = y_0, \tag{1.2.1.1}$$

where $y_0 \in \Omega \subset \mathbb{R}^d$, with $d \geq 1$, and $y : \ [0, T] \to \Omega$ is a vector function and $f : \ \Omega \times [0, T] \to \Omega$ is a given map such that (1.2.1.1) possesses a unique solution for all $y_0 \in \Omega$. Before we begin the analysis, we formulate some notations and definitions.

(i) Define the inner product:

$$x \cdot y = \sum_{i=1}^{d} x_i y_i, \quad x = (x_1, \ldots, x_d), \ y = (y_1, \ldots, y_d) \in \mathbb{R}^d.$$

(ii) $L^p([a,b];\Omega)$ is the space of p-integrable ($1 \le p < \infty$) or essentially bounded functions ($p = \infty$);

(iii) $H^k([a,b];\Omega)$ is the Sobolev space of functions with derivatives through order k in $L^2([a,b];\Omega)$ (k is a nonnegative integer);

(iv) $C([a,b];\Omega)$ denotes the space of continuous functions;

(v) $\mathcal{P}^{(m)}([a,b];\Omega)$ is the space of all polynomials of degree not exceeding m;

(vi) $||\cdot||_{\infty,n} := \sup\limits_{t\in \bar{I}_n} |\cdot|$, where $|\cdot|$ is the Euclidean norm; define, for nonnegative integer k,

$$||\cdot||_{k,n}^2 := \sum_{i=0}^{k} (\cdot^{(i)}, \cdot^{(i)})_n,$$

where $(\cdot,\cdot)_n$ denotes the inner product in $L^2(I_n;\Omega)$; define also

$$||\cdot||_\infty := \max\{||\cdot||_{\infty,n} : \; n = 1,\ldots,N\}; \; ||\cdot||_k^2 := \sum_{n=1}^{N} ||\cdot||_{k,n}^2.$$

Now we present the weak form of (1.2.1.1). On each interval I_n, form the inner product of (1.2.1.1) with v_n in $H^1(I_n;\Omega)$ and integrate by parts:

$$y(t_n)\cdot v_n(t_n) = y(t_{n-1})\cdot v_n(t_{n-1}) + \int_{I_n} [y\cdot v'_n + f(y)\cdot v_n]dt, \tag{1.2.1.2}$$

where $f(y)$ denotes the function $t \to f(t, y(t)) : \; [0,T] \to \Omega$. Then sum over all n equations (1.2.1.2), observing (1.2.1.1) to obtain the following variational equation:

$$\begin{aligned} & y(t_0)\cdot[V_0 - v_1(t_0)] + \sum_{n=1}^{N-1} y(t_n)\cdot[v_n(t_n) - v_{n+1}(t_n)] + y(t_N)\cdot v_N(t_N) \\ - & \sum_{n=1}^{N}\int_{I_n} y\cdot v'_n dt = y_0\cdot V_0 + \sum_{n=1}^{N}\int_{I_n} f(y)\cdot v_n dt, \end{aligned} \tag{1.2.1.3}$$

which is to hold for all

$$\tilde{v} := (V_0, v_1, \ldots, v_N) \in \mathcal{V} := \Omega \times \prod_{n=1}^{N} H^1(I_n;\Omega).$$

The space $\mathcal{V}$ will be endowed with the norm

$$\|\tilde{v}\|_{\mathcal{V}} := \left\{ |V_0|^2 + \sum_{n=1}^{N} \|v_n\|_{1,n}^2 \right\}^{1/2}.$$

This suggests the following variational problem: Find

$$\tilde{u} := (U_0, \ldots, U_N, u_1, \ldots, u_N) \in \mathcal{U} := \Omega^{N+1} \times \prod_{n=1}^{N} L^2(I_n; \Omega),$$

such that

$$\begin{aligned} &U_0 \cdot [V_0 - v_1(t_0)] + \sum_{n=1}^{N-1} U_n \cdot [v_n(t_n) - v_{n+1}(t_n)] + U_N \cdot v_N(t_N) \\ -& \sum_{n=1}^{N} \int_{I_n} u_n \cdot v'_n dt = y_0 \cdot V_0 + \sum_{n=1}^{N} \int_{I_n} f(u_n) \cdot v_n dt, \quad \forall \tilde{v} \in \mathcal{V}. \end{aligned} \tag{1.2.1.4}$$

Locally, the weak form (1.2.1.4) is equivalent to finding u_n in $L^2(I_n; \Omega)$ and U_n in Ω such that

$$\begin{cases} U_0 = y_0, \\ U_n \cdot v_n(t_n) = U_{n-1} \cdot v_n(t_{n-1}) + \displaystyle\int_{I_n} [u_n \cdot v'_n + f(u_n) \cdot v_n] dt, \end{cases} \tag{1.2.1.5}$$

for all $v_n \in H^1(I_n; \Omega)$ and $n = 1, \ldots, N$.

Theorem 1.2.1.1. [Delfour and Dubeau (1986)]

(a) *There exists a unique solution $\tilde{u} \in \mathcal{U}$ to the variational equation (1.2.1.4).*

(b) *Moreover,*

$$\tilde{u} = (y(t_0), \ldots, y(t_N), y|_{I_1}, \ldots, y|_{I_N}),$$

where y is the solution of problem (1.2.1.3) and $y|_{I_n}$ denotes the restriction of the function y to the interval I_n.

Now we introduce the Galerkin scheme corresponding to the weak form (1.2.1.5). Define the finite-dimensional subspaces $\mathcal{U}_h$ of $\mathcal{U}$, and $\mathcal{V}_h$ of $\mathcal{V}$ as follows:

$$\mathcal{U}_h = \left\{ \tilde{u}_h \;\middle|\; \begin{array}{l} \tilde{u}_h = (U_0^h, \ldots, U_N^h, u_1^h, \ldots, u_N^h) \in \mathcal{U} \text{ such that } u_n^h \in \mathcal{P}^{(m)}(I_n; \Omega) \\ \text{subject to } J\ (\geq 0) \text{ additional conditions for } n = 1, \ldots, N. \end{array} \right\},$$

$$\mathcal{V}_h = \left\{ \tilde{v}_h \;\middle|\; \begin{array}{l} \tilde{v}_h = (V_0^h, v_1^h, \dots, v_N^h) \in \mathcal{V} \text{ such that} \\ v_n^h \in \mathcal{P}^{(m+1-J)}(I_n;\Omega) \text{ for } n = 1, \dots, N. \end{array} \right\},$$

where m and J are nonnegative integers such that $m+1-J \geq 0$. Note also that

$$\dim \mathcal{U}_h = [1 + (m+2-J)N] \dim \Omega = \dim \mathcal{V}_h.$$

With the above definition, the approximation scheme for (1.2.1.5) is defined to find $\tilde{u}_h$ in $\mathcal{U}_h$ such that $U_0 = y_0$ and

$$\begin{cases} U_n^h \cdot v_n^h(t_n) - \displaystyle\int_{I_n} u_n^h \cdot v_n^{h\prime} dt = U_{n-1}^h \cdot v_n^h(t_{n-1}) + \int_{I_n} f(u_n^h) \cdot v_n^h dt, \\ J \text{ additional conditions on } u_n^h, \end{cases} \tag{1.2.1.6}$$

for all v_n^h in $\mathcal{P}^{(m+1-J)}(I_n;\Omega)$ and $n = 1, \dots, N$. Delfour and Dubeau [40] showed that (1.2.1.6) possesses a unique solution whenever h is small enough.

Remark 1.2.1.1.

(i) For $J = 0$ we obtain the completely discontinuous Galerkin methods;

(ii) For $0 < J \leq m+1$, and on each interval I_n, the J conditions are of the form

$$u_n^h(t_{n_l}) = U_{n_l}^h, \quad l = 1, \dots, J, \tag{1.2.1.7}$$

where $n_l \in \{0, \dots, N\}$. These Galerkin methods will be referred to as nodal methods:

(1) for $J = 1$, i.e., $u_n^h(t_n) = U_n^h$, $n = 1, \dots, N$, the nodal method coincides with the DG scheme of Lesaint and Raviart [82].

(2) for $J = 2$, i.e., $u_n^h(t_{n-1}) = u_{n-1}^h(t_{n-1}) = U_{n-1}^h$, the nodal methods become the continuous Galerkin methods of Hulme [73] [74].

(3) for $J = m+1$ in the nodal methods, we obtain multistep methods (see e.g., Butcher [30]).

(4) for $J < m+1$, the nodal methods reduce to hybrid methods (see e.g., Gear [50]).

(iii) for $J = 1$, on each interval I_n,

$$\alpha_n u_n^h(t_n) + (1-\alpha_n)u_{n+1}^h(t_n) = U_n^h,$$

the method is called the α-method (see Delfour, Hager and Trochu [41]).

1.2.2 Superconvergence

Now we establish the convergence results in two main theorems. The first theorem shows that if the solution of (1.2.1.1) belongs to $H^{m+1}([0,T];\Omega)$, the L^2 and nodal errors are proportional to h^{m+1}. The second theorem states that under appropriate assumption on the function f there is an asymptotic superconvergence at the mesh points proportional to h^{2m+2-J} $(0 \leq J \leq m+1)$. We assume that h is sufficiently small, in order to guarantee the existence of a unique solution $\tilde{u}$ to (1.2.1.6). C will denote a generic constant independent of h.

Theorem 1.2.2.1. [L^2 **and Nodal Errors]**

Assume that the solution y of (1.2.1.1) belongs to $H^{m+1}([0,T];\Omega)$. For $M > 1$, assume that on the first $M-1$ intervals the solution of (1.2.1.6) is such that

$$\max\{|U_n^h - y(t_n)| : \; n = 0, \ldots, M-1\} \leq Ch^{m+1}||y^{(m+1)}||_0,$$

and for $j = 0, \ldots, m+1$,

$$\left\{ \sum_{n=1}^{M-1} ||u_n^h - y||_{j,n}^2 \right\}^{1/2} \leq Ch^{m+1-j}||y^{(m+1)}||_0.$$

Thus,

$$\max\{|U_n^h - y(t_n)| : \; n = 0, \ldots, N\} \leq Ch^{m+1}||y^{(m+1)}||_0,$$

and for $j = 0, \dots, m+1$,

$$||u^h - y||_j \le Ch^{m+1-j}||y^{(m+1)}||_0,$$

where $u^h = \sum_{n=1}^N u_n \chi_{I_n}$ *and* $||\cdot||_j := \left\{\sum_{n=1}^N ||\cdot||_{j,n}^2\right\}^{1/2}$. *Here* χ_{I_n} *is the characteristic function of* I_n.

Proof. See Delfour and Dubeau [40]. It can also be found in Chapter 2 as a special case of Theorem 2.6.1.1.

Theorem 1.2.2.2. [Superconvergence]

Assume that the assumptions of Theorem 1.2.2.1 hold. Assume also that

(i) the matrix

$$A(t) := (a_{i,j}(t))_{i,j=1}^d,$$

with $a_{i,j} := \dfrac{\partial f_i}{\partial y_j}(t, y)$, *exists and that its columns belong to* $H^{m+1-J}([0,T];\Omega)$, *and*

(ii) there exist a neighborhood V *of the origin* y *in* Ω *and a positive constant* B *such that*

$$|f(t,x) - f(t,y) - A(t)(x-y)| \le B|y-x|^2,$$

for all t *and all* x *in* $y + V$. *Then,*

$$\begin{aligned} \max\{|U_n^h - y(t_n)| : \ n = 0, \dots, N\} &\le C||u^h - y||_0[h^{m+1-J} + ||u^h - y||_0] \\ &\le Ch^{2m+2-J} \end{aligned}$$

Proof. See Delfour and Dubeau [40]. It will also be derived in Chapter 2 as a special case of Theorem 2.6.1.2.

1.3 The discretized DG method for ODEs

The discussion in this section originates from Brunner [23], which is based on Lesaint and Raviart [82].

1.3.1 The comparison of the discretized DG method with the collocation method

We recall the DG time-stepping scheme for (1.1.1.1): For $n = 1, \cdots, N$, find $Y|_{I_n} \in \mathcal{P}^{(m)}(I_n)$, such that

$$\int_{I_n} (Y' - f(t,Y))X\,dt + Y^+_{n-1}X^+_{n-1} = Y^-_{n-1}X^+_{n-1}, \quad \forall X \in \mathcal{P}^{(m)}(I_n). \tag{1.3.1.1}$$

Here we set $Y^-_0 := y_0$. Suppose now that the integrals in (1.3.1.1) are approximated by interpolatory $(m+1)$-point quadrature formulas with abscissas $t_{n,j} := t_n + c_j h_n$ $(0 =: c_0 < c_1 < \ldots < c_m \leq 1)$ and weights w_j $(j = 0, 1, \ldots, m)$. We denote the resulting *discretized DG solution* in $\mathcal{V}_N^{(m)}$ by $\tilde{Y}$. The fully discretized version of (1.3.1.1) is then given by

$$h_n \sum_{j=0}^{m} w_j[\tilde{Y}'(t_{n,j}) - f(t_{n,j}, \tilde{Y}(t_{n,j}))]X(t_{n,j}) + \tilde{Y}(t_n^+)X(t_n^+) - \tilde{Y}(t_n^-)X(t_n^-) = 0, \tag{1.3.1.2}$$

for all $X \in \mathcal{P}^{(m)}(I_n)$. Let

$$\tilde{Y}_n := \tilde{Y}(t_n^-), \quad \tilde{Y}_{n,0} := \tilde{Y}(t_n^+)\ (= \tilde{Y}(t_{n,0}^+)), \quad \tilde{Y}_{n,j} := \tilde{Y}(t_{n,j})\ (j = 1, \ldots, m),$$

and let $L_j(v)$ be the jth Lagrange canonical polynomial (of degree $m-1$) corresponding to the points $\{c_i :\ i = 1, \ldots, m\}$. Moreover, denote by $\{X_j :\ j = 0, 1, \ldots, m\}$ a (canonical) basis for $\mathcal{P}^{(m)}(I_n)$ so that

$$X_i(t_n + c_j h_n) = \delta_{i,j} \quad (i,\ j = 0, 1, \ldots, m).$$

Since the restriction of $\tilde{Y}'$ to I_n is a polynomial of degree $m-1$ we may write

$$\tilde{Y}'(t_n + vh_n) = \sum_{j=1}^{m} L_j(v)\tilde{Y}'(t_{n,j}), \quad v \in (0,1],$$

and hence

$$\tilde{Y}(t_n + vh_n) = \tilde{Y}(t_n^+) + h_n \int_0^v \tilde{Y}'(t_n + sh_n)ds, \quad v \in (0,1]. \tag{1.3.1.3}$$

On the other hand, (1.3.1.2) with $X = X_0$ yields

$$h_n w_0[\tilde{Y}'(t_{n,0}) - f(t_{n,0}, \tilde{Y}(t_{n,0})] + \tilde{Y}(t_n^+) - \tilde{Y}(t_n^-) = 0,$$

implying that

$$\tilde{Y}(t_n^+) = \tilde{Y}_n + h_n w_0[f(t_{n,0}, \tilde{Y}(t_n^+)) - \sum_{j=1}^{m} L_j(c_0)\tilde{Y}'(t_{n,j})]. \tag{1.3.1.4}$$

For $X = X_i$ $(i = 1, \ldots, m)$, with $X_i(t_{n,j}) = \delta_{i,j}$, we obtain from (1.3.1.2) the equations

$$w_i[\tilde{Y}'(t_{n,i}) - f(t_{n,i}, \tilde{Y}(t_{n,i}))] = 0,$$

where $w_i \neq 0$. This result can be used in (1.3.1.4) to produce

$$\tilde{Y}(t_n^+) = \tilde{Y}_n + h_n w_0 f(t_{n,0}, \tilde{Y}(t_n^+)) - h_n \sum_{j=1}^{m} w_0 L_j(c_0) f(t_{n,j}, \tilde{Y}(t_{n,j})). \tag{1.3.1.5}$$

The identity (1.3.1.3) allows us to write

$$\tilde{Y}(t_{n,i}) = \tilde{Y}(t_n^+) + h_n \sum_{j=1}^{m} \beta_j(c_i) f(t_{n,j}, \tilde{Y}(t_{n,j})), \tag{1.3.1.6}$$

with

$$\beta_j(v) := \int_0^v L_j(s)ds \ (j = 1, \ldots, m),$$

and $\beta_j(c_i) =: a_{i,j}$. Hence, setting $\tilde{Y}_{n,i} := \tilde{Y}(t_{n,i})$ and recalling (1.3.1.5) we obtain

$$\tilde{Y}_{n,i} = \tilde{Y}_n + h_n w_0 f(t_{n,0}, \tilde{Y}(t_n^+)) + h_n \sum_{j=1}^{m} [a_{i,j} - w_0 L_j(c_0)] f(t_{n,j}, \tilde{Y}_{n,j}) \tag{1.3.1.7}$$

$(i = 1, \ldots, m)$. The equations (1.3.1.5) and (1.3.1.7) form a system of $m+1$ nonlinear algebraic equations for $\tilde{\mathbf{Y}}_n := (\ \tilde{Y}(t_n^+), \tilde{Y}_{n,1}, \ldots, \tilde{Y}_{n,m}\ )^T \in \mathbb{R}^{m+1}$: its form closely resembles the one corresponding to collocation at the points $\{t_{n,0}, t_{n,1}, \ldots, t_{n,m}\}$. We now show that these equations may indeed be interpreted as the stage equations of an implicit $(m+1)$-stage Runge-Kutta method. Let $b_j := \beta_j(1)$ $(j = 1, \ldots, m)$, and observe that

$$b_j = \int_0^1 L_j(s)ds = \sum_{k=0}^{m} w_k L_j(c_k) = w_0 L_j(c_0) + w_j,$$

because our interpolatory $(m+1)$-point quadrature formula is exact for polynomials of degree not exceeding m. This leads to the relationship

$$b_j - w_0 L_j(c_0) = w_j,$$

and hence by (1.3.1.7) to

$$\tilde{Y}_{n+1} := \tilde{Y}(t_{n+1}^-) = \tilde{Y}_n + h_n \sum_{j=0}^{m} w_j f(t_{n,j}, \tilde{Y}_{n,j}). \tag{1.3.1.8}$$

We conclude that (1.3.1.7) together with (1.3.1.5) and (1.3.1.8) represents a *collocation-based $(m+1)$-stage implicit Runge-Kutta method* for (1.1.1.1). We summarize the above discussion as the following theorem.

Theorem 1.3.1.1. *The fully discretized DG scheme (1.3.1.2) may lead to the collocation-based $(m+1)$-stage implicit Runge-Kutta method* $\{(1.3.1.7), (1.3.1.5), (1.3.1.8)\}$ *for (1.1.1.1).*

1.4 History of the DG methods for ODEs

In 1974, Lesaint and Raviart [82] gave the first analysis of the discontinuous Galerkin method when applied to ordinary differential equations. They showed that the method is strongly A-stable and has order $2m+1$ at the mesh points, and that the Gauss-Radau discretization of the DG method is also of order $2m+1$ when piecewise polynomials of degree m are used.

In 1981, Delfour, Hager, and Trochu [41] introduced a class of DG methods, the so called α-methods, for which they proved that the global L^2-convergence and nodal convergence rates are $m+1$ and $2m+1$. It is interesting that in 1986 Delfour and Dubeau [40] (refer to Remark 1.2.1.1 in Section 1.2.1) considered the discontinuous Galerkin method based on the mesh-dependent variational framework, which includes the "completely discontinuous" Galerkin methods, the α-methods, the continuous Galerkin methods, one-step methods of the Runge-Kutta type, hybrid and multi-step methods as special cases. It is shown that the convergence rate in the L^2-norm is $m+1$. The nodal-convergence rate can go up to $2m+2$, depending on the particular scheme under consideration.

In 1988, Johnson [77] gave an analysis of error control for the DG method for stiff ODEs and later in 1995, Estep [43] extended this analysis to general non-autonomous ODEs.

Recently Schötzau and Schwab [96] analyzed the *hp*-version of the discontinuous Galerkin methods. New a priori error bounds explicit in the time steps and in the approximation orders are derived and it is proven that the DG method gives spectral and exponential accuracy for problems with smooth and analytic time dependence, respectively. It is further shown that temporal singularities can be resolved at exponential rates of convergence if geometrically refined time steps are employed.

The readers may wish also to consult the 2000 survey paper [31] by Cockburn *et al.* for more applications of discontinuous Galerkin methods and for an extensive list of references.

Chapter 2

The discontinuous Galerkin method for VIDEs

The discontinuous Galerkin method for Volterra integral equations was first studied by Shaw and Whiteman [99] in 1996 extending the approach of [77] and [45]. In [99], they studied the discontinuous Galerkin method with a posteriori $L_p([0, t_i])$ error estimate for linear second-kind Volterra equations (compare [100]). Later in 1998 Larsson, Thomée, and Wahlbin [80] analyzed the discontinuous Galerkin method for linear parabolic integro-differential equations. Recently Brunner and Schözau [27] studied the *hp*-version of discontinuous Galerkin methods for parabolic Volterra integro-differential equations with weakly singular kernels.

2.1 The discontinuous Galerkin method for linear VIDEs

In this section, we consider the a priori error estimates, a posteriori estimates and superconvergence of DG method for linear Volterra integro-differential equations.

2.1.1 A priori error estimates for linear VIDEs

We study the scalar linear Volterra integro-differential equation,

$$\begin{cases} y'(t) + a(t)y(t) = \mathcal{V}(y)(t), \ t \in I = [0, T], \\ y(0) = y_0, \end{cases} \tag{2.1.1.1}$$

where $\mathcal{V}(y)(t) := \int_0^t k(t-s)y(s)ds$ and $a, \ k \in C[0, T]$.

We use the notations introduced in Section 1.1. We define the finite-dimensional space (cf. (1.1.1.2))

$$\mathcal{V}_N^{(m)} := \{\varphi \in L^2(I) : \ \varphi|_{I_n} \in \mathcal{P}^{(m)}(I_n), \ 1 \leq n \leq N\}, \tag{2.1.1.2}$$

where $\mathcal{P}^{(m)}(I_n)$ denotes the space of all (real) polynomials of degree not exceeding m. Then the DG method for (2.1.1.1) is : Find $Y \in \mathcal{V}_N^{(m)}$ such that

$$B_{DG}(Y, X) = Y_0^- X_0^+, \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.1.1.3}$$

where $Y_0^- = y_0$ and

$$\begin{aligned} B_{DG}(Y, X) \ &:= \ \sum_{n=1}^{M} \int_{I_n} (Y'(t) + a(t)Y(t) - \mathcal{V}(Y)(t)) \, X(t)dt \\ &+ \ \sum_{n=2}^{M} [Y]_{n-1} X_{n-1}^+ + Y_0^+ X_0^+. \end{aligned} \tag{2.1.1.4}$$

Note that the exact solution y of (2.1.1.1) satisfies

$$B_{DG}(y, X) = y_0 X_0^+, \quad \forall X \in \mathcal{V}_N^{(m)},$$

hence the Galerkin orthogonality property

$$B_{DG}(y - Y, X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.1.1.5}$$

holds true. We remark also that the DG method in (2.1.1.3) can be interpreted as a time-stepping scheme. For $n = 1, \cdots, N$, find $Y|_{I_n} \in \mathcal{P}^{(m)}(I_n)$, such that

$$\int_{I_n} (Y'(t) + a(t)Y(t) - \mathcal{V}(Y)(t))X dt + Y_{n-1}^+ X_{n-1}^+ = Y_{n-1}^- X_{n-1}^+, \tag{2.1.1.6}$$

for all $X \in \mathcal{P}^{(m)}(I_n)$. Here we set $Y_0^- := y_0$. We will refer this method as DG(m). When adopting the Picard iteration technique, we can easily prove that (2.1.1.3) has a unique solution.

Theorem 2.1.1.1. *Assume $\mathcal{A} := ||a||_I$, $\mathcal{B} := ||k||_I$. Then there is a constant C, independent of h_n, such that for $1 \le n \le M \le N$ the error of DG(m) for (2.1.1.1) satisfies*

$$||e||_{[0,t_M]} \le C \max_{n \le M} h_n^{m+1} ||y^{(m+1)}||_{\bar{I}_n},$$

with $C := C(t_M, \mathcal{A}, \mathcal{B})$ and $m = 0,\ 1$.

Proof. If $V \in \mathcal{V}_N^{(m)}$ is determined by $V_0^- = y_0$ and, for $1 \le n \le M \le N$, by

$$\int_{I_n} V' X dt + \int_{I_n} \{a(t)y(t) - \mathcal{V}(y)(t)\} X dt + V_{n-1}^+ X_{n-1}^+ = V_{n-1}^- X_{n-1}^+, \qquad (2.1.1.7)$$

for all $X \in \mathcal{V}_N^{(m)}$ $(m = 0,\ 1)$, then

$$||y - V||_{\bar{I}_n} \le C h_n^{m+1} ||y^{(m+1)}||_{\bar{I}_n}, \qquad (2.1.1.8)$$

We now compare Y to V. Setting $\mu := y - V$ *and* $\phi := Y - V \in \mathcal{V}_N^{(m)}$, we have $e = \mu - \phi$. We subtract (2.1.1.7) from (2.1.1.6) and get

$$\int_{I_n} \phi' X dt + \int_{I_n} -\{a(t)eX - \mathcal{V}(e)(t)X\} dt + \phi_{n-1}^+ X_{n-1}^+ = \phi_{n-1}^- X_{n-1}^+, \qquad (2.1.1.9)$$

for all $X \in \mathcal{V}_N^{(m)}$. We choose $X = \phi$ to obtain

$$\frac{1}{2}\left|\phi_{n-1}^+\right|^2 + \frac{1}{2}\left|\phi_n^-\right|^2 + \int_{I_n} -\{a(t)e\phi - \mathcal{V}(e)(t)\phi(t)\} dt = \phi_{n-1}^- \phi_{n-1}^+. \qquad (2.1.1.10)$$

Hence

$$\frac{1}{2}\left|\phi_n^-\right|^2 \le \frac{1}{2}\left|\phi_{n-1}^-\right|^2 + \int_{I_n} \left|\{a(t)e\phi - \mathcal{V}(e)(t)\phi(t)\}\right| dt. \qquad (2.1.1.11)$$

We now substitute $e = \mu - \phi$ into (2.1.1.11) and find that

$$\begin{aligned}
\left|\int_{I_n} a(t)e\phi dt\right| &= \left|\int_{I_n} a(t)(\phi-\mu)\phi dt\right| \\
&= \left|\int_{I_n} a(t)\phi^2 - \int_{I_n} a(t)\mu\phi dt\right| \\
&\leq \mathcal{A}\int_{I_n}\phi^2 dt + \frac{\mathcal{A}}{2}\int_{I_n}(\mu^2+\phi^2)dt \\
&= \frac{\mathcal{A}}{2}\int_{I_n}\mu^2 dt + \frac{3\mathcal{A}}{2}\int_{I_n}\phi^2 dt. \qquad (2.1.1.12)
\end{aligned}$$

$$\begin{aligned}
\left|\int_{I_n}\mathcal{V}(e)(t)\phi(t)dt\right| &\leq \mathcal{B}\int_{I_n}\phi^2(t)dt + \frac{1}{\mathcal{B}}\int_{I_n}\left(\mathcal{V}(e)(t_{n-1})\right)^2 dt \\
&+ \frac{1}{\mathcal{B}}\int_{I_n}\left(\int_{t_{n-1}}^{t} k(t-s)e(s)ds\right)^2 dt \\
&\leq \mathcal{B}\int_{I_n}\phi^2(t)dt + \frac{1}{\mathcal{B}}\int_{I_n}\left(\int_{I_n} k^2(t)dt\right)\cdot\left(\int_{I_n} e^2 dt\right)dt \\
&+ \frac{1}{\mathcal{B}}\int_{I_n}\left(\int_0^{t_{n-1}} k^2(t-s)ds\cdot\int_0^{t_{n-1}} e^2(s)ds\right)dt \\
&\leq \mathcal{B}\int_{I_n}\phi^2(t)dt + \mathcal{B}h_n^2\int_{I_n}e^2(t)dt + \mathcal{B}t_{n-1}h_n\int_0^{t_{n-1}}e^2(t)dt \\
&\leq (\mathcal{B}+2\mathcal{B}h_n^2)\int_{I_n}\phi^2(t)dt + 2\mathcal{B}h_n^2\int_{I_n}\mu^2 dt \\
&+ \mathcal{B}t_{n-1}h_n\int_0^{t_{n-1}}e^2(t)dt. \qquad (2.1.1.13)
\end{aligned}$$

Combining (2.1.1.11), (2.1.1.12), and (2.1.1.13), we obtain

$$\begin{aligned}
\left|\phi_n^-\right|^2 &\leq \left|\phi_{n-1}^-\right|^2 + (\mathcal{A}+4\mathcal{B}h_n^2)\int_{I_n}\mu^2 dt \\
&+ (3\mathcal{A}+4\mathcal{B}h_n^2+2\mathcal{B})\int_{I_n}\phi^2 dt + 2\mathcal{B}t_{n-1}h_n\int_0^{t_{n-1}}e^2(t)dt. \qquad (2.1.1.14)
\end{aligned}$$

Next we choose $X = (t - t_{n-1})\phi'$ in (2.1.1.9) to obtain

$$\begin{aligned}
\int_{I_n} (t - t_{n-1})(\phi')^2 dt &\leq \int_{I_n} (a(t))^2 e^2 \cdot (t - t_{n-1}) dt + \frac{1}{4}\int_{I_n} (t - t_{n-1})(\phi')^2 dt \\
&+ \int_{I_n} (\mathcal{V}(e)(t))^2 (t - t_{n-1}) dt + \frac{1}{4}\int_{I_n} (t - t_{n-1})(\phi')^2 dt \\
&\leq \mathcal{A}^2 h_n \int_{I_n} e^2 dt + \frac{1}{2}\int_{I_n} (t - t_{n-1})(\phi')^2 dt \\
&+ 2h_n^3 \mathcal{B}^2 \int_{I_n} e^2 dt + 2h_n^2 \mathcal{B}^2 t_{n-1} \int_0^{t_{n-1}} e^2 dt \\
&\leq 2\left(\mathcal{A}^2 h_n + 2h_n^3 \mathcal{B}^2\right) \int_{I_n} \left(\mu^2 + \phi^2\right) dt \\
&+ \frac{1}{2}\int_{I_n} (t - t_{n-1})(\phi')^2 dt + 2h_n^2 \mathcal{B}^2 t_{n-1} \int_0^{t_{n-1}} e^2 dt.
\end{aligned}$$

Since $\int_{I_n} |\phi|^2 \, dt \leq 2h_n \left|\phi_n^-\right|^2 + \frac{2}{3}h_n^3 ||\phi'||_{I_n}^2$, we find that

$$\int_{I_n} |\phi|^2 \, dt \leq 4h_n \left|\phi_n^-\right|^2 + \int_{I_n} \mu^2 dt + \frac{32}{3} h_n^3 \mathcal{B}^2 t_{n-1} \int_0^{t_{n-1}} e^2 dt, \tag{2.1.1.15}$$

provided that $\frac{16}{3}h_n^2 \left(\mathcal{A}^2 + 2h_n^2\mathcal{B}^2\right) \leq \frac{1}{2}$.

By combining (2.1.1.14) and (2.1.1.15), we see that

$$\begin{aligned}
\left|\phi_n^-\right|^2 &\leq 2\left|\phi_{n-1}^-\right| + (8\mathcal{A} + 4\mathcal{B}) \int_{I_n} \mu^2 dt \\
&+ 2(2\mathcal{B} t_{n-1} h_n + \frac{4}{3} h_n^2 \mathcal{B}^2 t_{n-1}) \int_0^{t_{n-1}} e^2 dt,
\end{aligned} \tag{2.1.1.16}$$

provided that $4h_n(3\mathcal{A} + 4\mathcal{B}h_n^2 + 2\mathcal{B}) \leq \frac{1}{2}$. It thus follows that

$$\left|\phi_M^-\right|^2 \leq C\left(\int_0^{t_{M-1}} \mathcal{B} \cdot e^2 dt + [(\mathcal{A} + \mathcal{B})t_M] ||\mu||_{[0,t_M]}^2\right). \tag{2.1.1.17}$$

The estimate of Theorem 2.1.1.1 is obtained from (2.1.1.17) directly when $m = 0$. For $m = 1$, we use (2.1.1.10) and the fact that $\phi \in \mathcal{V}_N^{(m)}$ implies that

$$||\phi||_{I_n}^2 \leq \left|\phi_{n-1}^+\right|^2 + \left|\phi_n^-\right|^2,$$

and this yields

$$\begin{aligned}\frac{1}{4}||\phi||^2_{I_n} &\leq |\phi^-_{n-1}|^2 + \frac{\mathcal{A}+2h_n^2\mathcal{B}}{2}\int_{I_n}\mu^2 dt + \frac{(3\mathcal{A}+2h_n^2\mathcal{B}+2\mathcal{B})}{2}\int_{I_n}\phi^2 dt\\ &+ \frac{1}{2}\mathcal{B}^2 t_{n-1}h_n\int_0^{t_{n-1}} e^2 dt.\end{aligned}$$

An analogous argument now leads to the estimate when $m = 1$.

2.1.2 A posteriori error estimates for linear VIDEs

We analyze the a posteriori error bound for DG(m) approximation to (2.1.1.1), by using the stability of the continuous dual problem associated with (2.1.1.1).

Theorem 2.1.2.1. *Assume that $\mathcal{A} := ||a||_I$, $\mathcal{B} := ||k||_I$, and let $m = 0,\ 1$. Then the DG(m) finite element solution Y for (2.1.1.1) satisfies, for $0 \leq M \leq N$,*

$$|y(t_M) - Y^-_M| \leq C||h_n^{m+1}R(Y)||_{[0,t_M]},$$

where $C := C(t_M, \mathcal{A}, \mathcal{B})$ and $R(Y) := \frac{|[Y]_{n-1}|}{h_n} + |a(t)Y - \mathcal{V}(Y)(t)|$ $(t \in I_n)$.

Proof. We study the dual problem of (2.1.1.1),

$$\begin{cases} -z' + a(t)z = \mathcal{V}^*(z)(t), & t \in (0, t_M),\\ z(t_M) = e^-_M, \end{cases} \tag{2.1.2.1}$$

where $\mathcal{V}^*(z)(t) := \displaystyle\int_t^{t_M} k(s-t)z(s)ds$. From the definition of B_{DG} in (2.1.1.4), we find that for all piecewise continuous functions, $x,\ z \in C(I)$, the exact solution of (2.1.2.1) satisfies

$$B_{DG}(x, z) = (x^-_M, e^-_M). \tag{2.1.2.2}$$

If we choose $x = e$ in (2.1.2.2) we obtain,

$$|e^-_M|^2 = B_{DG}(e, z), \tag{2.1.2.3}$$

and we know from the Galerkin orthogonality (2.1.1.5) that

$$|e_M^-|^2 = B_{DG}(e, z - X), \quad \forall X \in \mathcal{V}_N^{(m)}. \tag{2.1.2.4}$$

Now we define for $X \in \mathcal{V}_N^{(m)}$ $(m = 0,\ 1)$ and $1 \le n \le M \le N$,

$$\Re_m(X, M; n) := S(M) \left(h_n^m \, |[X]_{n-1}| + h_n^{m+1} \, ||a(t)X - \mathcal{V}(X)(t)|\,|_{\bar{I}_n} \right) \tag{2.1.2.5}$$

where $S(M) = \displaystyle\int_0^{t_M} \left| z^{(m+1)} \right| dt$, and z is the exact solution of (2.1.2.1).

We now show that

$$\left| e_M^- \right|^2 = |B_{DG}(e, z - X)| \le \max_{n \le M} \Re_m(Y, M; n). \tag{2.1.2.6}$$

Because of (2.1.1.1), and since $[y]_n = 0$ for all n, $Y_0^- = y_0$, and by the definition of B_{DG} in (2.1.1.3), we have

$$\begin{aligned} B_{DG}(e, z - X) \;&=\; -\sum_{n=1}^{M} \int_{I_n} \{ (Y' + a(t)Y)\,(z - X) \\ &-\; (\mathcal{V}(Y)(t))\,(z - X) \} dt - \sum_{n=1}^{M} [Y]_{n-1} (z - X)_{n-1}^{+}. \end{aligned}$$

Hence,

$$\begin{aligned} |B_{DG}(e, z - X)| \;&\le\; \left| \sum_{n=1}^{M} \int_{I_n} Y'(z - X) dt \right| \\ &+\; \left| \sum_{n=1}^{M} \int_{I_n} (a(t)Y - \mathcal{V}(Y)(t))\,(z - X) dt \right| \\ &+\; \left| \sum_{n=1}^{M} [Y]_{n-1} (z - X)_{n-1}^{+} \right| \\ &=:\; I + II + III. \end{aligned}$$

We set $X := P(z)$, where $P(z)$ denotes the projection of z onto $V_N^{(m)}$ and satisfies,

$$X_n^- = z(t_n) \text{ for } m = 0,\ 1; \text{ and } \int_{I_n} (X - z) dt = 0 \text{ for } m = 1. \tag{2.1.2.7}$$

Furthermore, we know that the estimates

$$\|z - P_0(z)\|_{\bar{I}_n} \leq h_n^m \int_{I_n} |z^{(m+1)}|\, dt \text{ and } \|z - P_0(z)\|_{\bar{I}_n} \leq \max_{n\leq M} h_n^{(m+1)} ||z^{(m+1)}||_{\bar{I}_n}, \tag{2.1.2.8}$$

hold for $m = 0,\ 1$. Thus, $I = 0$ and

$$\begin{aligned} II &\leq \sum_{n=1}^{M} h_n \|a(t)Y - \mathcal{V}(Y)(t)\|_{\bar{I}_n} \|z - P(z)\|_{\bar{I}_n} \\ &\leq \max_{n\leq M} h_n \,||a(t)Y - \mathcal{V}(Y)(t)||\,|_{\bar{I}_n} \sum_{n=1}^{M} h_n^m \int_{I_n} |z^{(m+1)}|\, dt \\ &=: S(M) \max_{n\leq M} h_n^{(m+1)} \,||a(t)Y - \mathcal{V}(Y)(t)||\,|_{\bar{I}_n}. \end{aligned}$$

The same kind of argument shows for $m = 0$ and 1,

$$III \leq S(M) \max_{n\leq M} h_n^m \,|[Y]_{n-1}|\,.$$

Hence, (2.1.2.6) holds true.

Now we prove the stability of (2.1.2.1).

Lemma 2.1.2.1. *If $\mathcal{A} := ||a||_I$ and $\mathcal{B} := ||k||_I$, then the solution z of (2.1.2.1) satisfies*

$$S(M) := \int_0^{t_M} |z^{(m+1)}|\, dt \leq C\, |e_M^-|\,,$$

where $C := C(t_M, \mathcal{A}, \mathcal{B})$.

Proof. Taking $t = t_M - s$ in (2.1.2.1) and setting $\psi(s) = z(t_M - s)$, (2.1.2.1) can be rewritten as

$$\begin{cases} \psi'(s) + a(t_M - s)\psi(s) = \displaystyle\int_0^s k(t-\nu)\psi(\nu)d\nu \\ \psi(0) = e_M^-. \end{cases} \tag{2.1.2.9}$$

Dirichlet's formula applied to (2.1.2.9) yields

$$\psi(s) = \psi(0) + \int_0^s Q(s,\nu)\psi(\nu)d\nu, \tag{2.1.2.10}$$

where $Q(s,\nu) := a(t_M - \nu) + \int_\nu^s k(\tau - \nu)d\tau$. Thus,

$$|\psi(s)| \leq |\psi(0)| + \int_0^s |Q(s,\nu)|\, |\psi(\nu)|\, d\nu.$$

It follows from the well-known Gronwall's lemma [26] that

$$|\psi(s)| \leq |\psi(0)| + \int_0^s |Q(s,\nu)| \exp\left(\int_\nu^s |Q(\tau,\tau)| d\tau\right) |\psi(0)| d\nu,$$

and so

$$\begin{aligned}
\int_0^{t_M} |z'(t)|\, dt &= \int_0^{t_M} |\psi'(s)|\, ds \\
&\leq \int_0^{t_M} \left\{ |a(t_M - s)|\, |\psi(s)| + \int_0^s |k(\nu - s)|\, |\psi(\nu)|\, d\nu \right\} ds \\
&\leq |\psi(0)| \int_0^{t_M} \{ |a(t_M - s)| \left(1 + \int_0^s |Q(s,\nu)| \exp\left(\int_\nu^s |Q(\tau,\tau)| d\tau\right) d\nu\right) \\
&+ \int_0^s |k(\nu - s)| \left(1 + \int_0^\nu |Q(\nu,\mu)| \exp\left(\int_\mu^\nu |Q(\tau,\tau)|\, d\tau\right) d\mu\right) d\nu \} ds \\
&= C|\psi(0)|,
\end{aligned}$$

where $C := C(t_M, \mathcal{A}, \mathcal{B})$. The similar argument can lead to

$$\int_0^{t_M} |z''(t)| dt \leq C|\psi(0)|.$$

The proof of the lemma has been completed.

From (2.1.2.6), (2.1.2.7), (2.1.2.8), and Lemma 2.1.2.1, we obtain

$$\begin{aligned}
|e_M^-| &\leq \frac{S(M)}{|e_M^-|} \max_{n \leq M} \{ h_n^m \, |[Y]_{n-1}| + h_n^{m+1} \|Y(t) - \mathcal{V}(Y)(t)\|_{\bar{I}_n} \} \\
&\leq C h_n^{m+1} \max_{n \leq M} \left\{ \frac{|[Y]_{n-1}|}{h_n} + \|a(t)Y - \mathcal{V}(Y)(t)\|_{\bar{I}_n} \right\}. \qquad (2.1.2.11)
\end{aligned}$$

This concludes the proof of Theorem 2.1.2.1.

2.1.3 Efficiency of the a posteriori error estimator

If its upper bound is large, the a posteriori error estimator $||h_n^{m+1}R(Y)||_{[0,t_M]}$ in Theorem 2.1.2.1 cannot efficiently indicate the error. Therefore we need to derive a sharper upper bound of the a posteriori error estimate. This estimate is called "efficiency of a posteriori error estimator" (see also Ainsworth and Oden [2]).

Theorem 2.1.3.1. *Under the assumptions of Theorem 2.1.2.1 and Theorem 2.1.1.1, we have*

$$\left|y(t_M) - Y_M^-\right| \leq C||h_n^{m+1}R(Y)||_{[0,t_M]} \leq Ch^{m+1}\left(||y'(t)||_{[0,t_M]} + ||y^{(m+1)}(t)||_{[0,t_M]}\right)$$

where $C := C(t_M, \mathcal{A}, \mathcal{B})$ is independent of the mesh size h.

Proof. We only need to bound the term $||h_n^{m+1}R(Y)||_{[0,t_M]}$ in Theorem 2.1.2.1:

$$\begin{aligned}
||h_n^{m+1}R(Y)||_{[0,t_M]} &\leq ||h_n^m[Y]_{n-1}||_{[0,t_M]} + h_n^{m+1}||a(t)Y - \mathcal{V}(Y)(t)||_{[0,t_M]} \\
&= ||h_n^m(Y_{n-1}^+ - y + y - Y_{n-1}^-)||_{[0,t_M]} + ||h_n^{m+1}(-y'(t) - a(t)y(t) \\
&+ \mathcal{V}(y)(t) + a(t)Y - \mathcal{V}(Y)(t))||_{[0,t_M]} \\
&\leq 2||h_n^m e||_{[0,t_M]} + ||h_n^{m+1}y'(t)||_{[0,t_M]} \\
&+ ||h_n^{m+1}(-a(t)e(t) + \mathcal{V}(e)(t))||_{[0,t_M]} \\
&\leq 2||h_n^m e||_{[0,t_M]} + ||h_n^{m+1}y'(t)||_{[0,t_M]} \\
&+ \mathcal{A}||h_n^{m+1}e||_{[0,t_M]} + t_M\mathcal{B}||h_n^{m+1}e||_{[0,t_M]}.
\end{aligned}$$

Then we appeal to Theorem 2.1.1.1 to complete the proof.

2.2 The discontinuous Galerkin method for semilinear VIDEs

We now extend a priori and a posteriori error estimates of Section 2.1 to semilinear VIDEs.

2.2.1 A priori error estimates for semilinear VIDEs

We study the scalar semilinear Volterra integro-differential equation

$$\begin{cases} y'(t) + a(t)y(t) = \mathcal{V}_G(y)(t), & t \in I = [0,T] \\ y(0) = y_0, \end{cases} \tag{2.2.1.1}$$

where $\mathcal{V}_G(y)(t) := \int_0^t k(t-s)G(y(s))ds$, and $a,\ k \in C(I)$. Furthermore assume G satisfies

$$|G(y_1) - G(y_2)| \leq L|y_1 - y_2|, \tag{2.2.1.2}$$

for all $y_1,\ y_2 \in \Omega \subset \mathbb{R}$. We begin with the definition of the DG(m) scheme to (2.2.1.1): Find $Y \in \mathcal{V}_N^{(m)}$ such that

$$B_{DG}(Y,X) = F_{DG}(X), \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.2.1.3}$$

where

$$\begin{aligned} B_{DG}(Y,X) &:= \sum_{n=1}^{M} \int_{I_n} \{Y'(t)X(t) + a(t)Y(t)X(t) \\ &- (\mathcal{V}_G(Y)(t))\, X(t)\}dt \\ &+ \sum_{n=1}^{M-1} [Y]_n X_n^+ + Y_0^+ X_0^+, \qquad (2.2.1.4) \\ F_{DG}(X) &:= y_0 X_0^+. \qquad (2.2.1.5) \end{aligned}$$

We note that

$$B_{DG}(Y,X) - B_{DG}(y,X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}. \tag{2.2.1.6}$$

The DG method (2.2.1.3) can again be interpreted as a time-stepping scheme: Find $Y|_{I_n} \in \mathcal{P}^{(m)}(I_n)$, $n = 1, \cdots, M$, such that,

$$\begin{aligned} &\int_{I_n} \{Y'(t)X(t) + a(t)Y(t)X(t) - \left(\int_{t_{n-1}}^{t} k(t-s)G(Y)ds\right) X(t)\}dt + Y_{n-1}^{+}X_{n-1}^{+} \\ = \; & Y_{n-1}^{-}X_{n-1}^{+} + \int_{I_n} \left(\int_0^{t_{n-1}} k(t-s)G(Y)ds\right) X(t)dt \end{aligned} \qquad (2.2.1.7)$$

Theorem 2.2.1.1. *Suppose that* $\mathcal{A} := ||a||_I$, $\mathcal{B} := ||k||_I$. *Then there is a constant* C, *independent of* h_n, *such that for* $1 \le n \le M \le N$, *the error of DG(m) for (2.2.1.1) satisfies*

$$||e||_{[0,t_M]} \le C \max_{n \le M} h_n^{m+1} ||y^{(m+1)}||_{\bar{I}_n},$$

with $C := C(t_M, L, \mathcal{A}, \mathcal{B})$ *and* $m = 0,\ 1$.

Proof. If $V \in \mathcal{V}_N^{(m)}$ is determined by $V_0^- = y_0$ and by

$$\int_{I_n} V'Xdt + \int_{I_n} \{a(t)y(t) - \mathcal{V}_G(Y)(t)\}Xdt + V_{n-1}^{+}X_{n-1}^{+} = V_{n-1}^{-}X_{n-1}^{+}, \qquad (2.2.1.8)$$

for all $X \in \mathcal{V}_N^{(m)}$ $(m = 0,\ 1)$ and for $1 \le n \le M \le N$, then it follows from the definition of the interpolant (2.1.1.8) that

$$||y - V||_{\bar{I}_n} \le Ch_n^{m+1}||y^{(m+1)}||_{\bar{I}_n}. \qquad (2.2.1.9)$$

Setting $\mu := y - V$ and $\phi := Y - V \in \mathcal{V}_N^{(m)}$, we have $e = \mu - \phi$. We subtract (2.2.1.8) from (2.2.1.3) and obtain

$$\int_{I_n} \phi' X dt + \int_{I_n} -\{a(t)eX - \int_0^t k(t-s)\left(G(Y) - G(y)\right) dsX\}dt + \phi_{n-1}^{+}X_{n-1}^{+} = \phi_{n-1}^{-}X_{n-1}^{+},$$

for all $X \in \mathcal{V}_N^{(m)}$. The remaining steps follow exactly those of Theorem 2.1.1.1 after we use the Lipschitz condition (2.6.1.2).

2.2.2 A posteriori error estimates for semilinear VIDEs

In this section, we derive the a posteriori estimates of the DG(m) in the mesh-point sense and the general a global posteriori error estimates of DG(0) for (2.2.1.1).

Theorem 2.2.2.1. *Assume that $\mathcal{A} := ||a||_I$, $\mathcal{B} := ||k||_I$, and the function G satisfies*

$$|G_y(u)| \leq L, \quad \forall u \in \Omega.$$

Then the error of the DG(m) approximation to (2.2.1.1) satisfies

$$\left|y(t_M) - Y_M^-\right| \leq C||h_n^{m+1}R(Y)||_{[0,t_M]},$$

with $m = 0$ and 1, $C := C(t_M, \mathcal{A}, \mathcal{B}, L)$ and

$$R(Y) = \frac{|[Y]_{n-1}|}{h_n} + |a(t)Y - \int_0^t k(t-s)G(Y(s))ds|.$$

Proof. Recall (2.2.1.6):

$$B_{DG}(Y, X) - B_{DG}(y, X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}.$$

We write this as

$$\tilde{D}(e, X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.2.2.1}$$

where

$$\begin{aligned} \tilde{D}(W, X) \quad := \quad & \sum_{n=1}^{M} \int_{I_n} \{W'X + a(t)WX \\ - \quad & \int_0^t k(t-s) \int_0^1 G_1\left(ry + (1-r)Y\right) dr W(s) \\ & ds X(t)\} dt \\ + \quad & \sum_{n=2}^{M} [W]_{n-1} X_{n-1}^+ + W_0^+ X_0^+. \end{aligned} \tag{2.2.2.2}$$

We consider the linearized dual problem of (2.2.1.1):

$$\begin{cases} -z' + a(t)z = \displaystyle\int_t^{t_M} k(s-t)\tilde{A}(s)z(s)ds, \quad t_M > t > 0, \\ z(t_M) = e_M^-, \end{cases} \tag{2.2.2.3}$$

where $\tilde{A}(s) := \displaystyle\int_0^1 G_1\left(ry + (1-r)Y\right) dr$. We note that, for any piecewise continuous function x,

$$\tilde{D}(x, z) = x_M^- e_M^-, \tag{2.2.2.4}$$

Selecting $x = e$ in (2.2.2.4), we have

$$[e_M^-]^2 = \tilde{D}(e, z). \tag{2.2.2.5}$$

In view of (2.2.2.1), we obtain

$$[e_M^-]^2 = \tilde{D}(e, z - X), \quad \forall X \in \mathcal{V}_N^{(m)}, \quad m = 0,\ 1.$$

Similarly to the proof of Theorem 2.1.2.1, we can continue the analysis, to find

$$\begin{aligned} \tilde{D}(e, z - X) \ &:= \ -\sum_{n=1}^{M} \int_{I_n} \{(Y' + a(t)Y)(z - X) - (\mathcal{V}_G(Y)(t)) \cdot \\ &\quad (z - X)\}dt - \sum_{n=1}^{M} [Y]_{n-1}(z - X)_{n-1}^{+}. \end{aligned}$$

Hence,

$$\begin{aligned} \left|\tilde{D}(e, z - X)\right| \ &\leq \ \left|\sum_{n=1}^{M} \int_{I_n} Y'(z - X)dt\right| \\ &+ \ \left|\sum_{n=1}^{M} \int_{I_n} (a(t)Y - \mathcal{V}_G(Y)(t)) \cdot (z - X)dt\right| \\ &+ \ \left|\sum_{n=1}^{M} [Y]_{n-1}(z - X)_{n-1}^{+}\right| \\ &=: \ I + II + III \\ &\leq \ S(M) \max_{n \leq M} \{h_n^m [Y]_{n-1} \\ &+ \ h_n^{m+1} \left|a(t)Y - \mathcal{V}_G(Y)(t)\right|\}, \end{aligned} \tag{2.2.2.6}$$

where $S(M)$ can be easily estimated as in Lemma 2.1.2.1,

$$S_1(M) := \int_0^{t_M} \left|z^{(m+1)}\right| dt \le C \left|e_M^-\right|. \tag{2.2.2.7}$$

Here $C := C(t_M, \mathcal{A}, \mathcal{B}, L)$. Combining (2.2.2.5), (2.2.2.6), and (2.2.2.7), we finish the proof.

Theorem 2.2.2.2. *Assume that $\mathcal{A} := ||a||_I$, $\tilde{\mathcal{B}} := \int_0^{t_M} |k(t)|dt$, and G satisfies*

$$\ell|u-v| \le |G(u)-G(v)| \le L|u-v|; \ |G_1(u)-G_1(v)| \le \tilde{L}|u-v|, \tag{2.2.2.8}$$

for all u, $v \in \Omega$. Then the error of the DG(0) approximation to (2.2.1.1) satisfies

$$\begin{aligned} ||e||^2_{[0,t_M]} &\le \max_{n\le M} C\left(|[Y]_{n-1}| + h_n\left|a(t)Y - \mathcal{V}_G(Y)(t)\right|\right)^2 \\ &+ C\left([Y]_{M-1}\right)^2 + C\left(a(t)Y + \mathcal{V}_G(Y)(t)\right)^2. \end{aligned}$$

Proof. We begin the proof with the related linearized form of (2.2.2.2),

$$\begin{aligned} D(W,X) &:= \sum_{n=1}^{M} \int_{I_n} \left\{ W'X + a(t)WX - \int_0^t k(t-s)A(s)W(s)dsX(t) \right\} dt \\ &+ \sum_{n=2}^{M} [W]_{n-1}X_{n-1}^+ + W_0^+ X_0^+ \\ &= \sum_{n=1}^{M} \int_{I_n} \left\{ -WX' + a(t)WX - \left(\int_0^t k(t-s)A(s)W(s)ds\right)X(t) \right\} dt \\ &+ \sum_{n=2}^{M} W_{n-1}^-[X]_{n-1} + W_n^- X_n^-, \end{aligned} \tag{2.2.2.9}$$

where $A(s) := G_1(y(s))$.

We consider the linearized dual problem of (2.2.1.1):

$$\begin{cases} -z' + a(t)z = \int_t^{t_M} k(s-t)A(s)z(s)ds, \quad t_M > t > 0, \\ z(t_M) = e_M^-. \end{cases} \tag{2.2.2.10}$$

We note that Z, the DG(m) approximation to z, solves

$$D(X, Z) = X_M^- e_M^-. \tag{2.2.2.11}$$

We claim also, for any piecewise continuous function x,

$$D(x, z) = x_M^- e_M^-. \tag{2.2.2.12}$$

The choice $x = e$ in (2.2.2.12) gives

$$D(e, z) = |e_M^-|^2 = \tilde{D}(e, z) + (D - \tilde{D})(e, z).$$

This, together with (2.2.2.1), yields the error representation formula

$$|e_M^-|^2 = \tilde{D}(e, z - X) + (D - \tilde{D})(e, z), \quad \forall X \in \mathcal{V}_N^{(m)}. \tag{2.2.2.13}$$

Now we define for $X \in \mathcal{V}_N^{(m)}$ and $1 \leq n \leq M \leq N$,

$$\begin{aligned} &\Re_m(X, M; n) \\ := \;& S(M) \left(|[X]_{n-1}| + h_n \left| a(t)X - \mathcal{V}_G(X)(t) \right| \right), \end{aligned} \tag{2.2.2.14}$$

for $m = 0$ and 1.

Note that by the assumption (2.2.2.8),

$$\begin{aligned} \left|(D - \tilde{D})(e, z)\right| &= \left| \int_0^{t_M} \int_0^t k(t-s)A(s)e(s)dsz(t)dt \right. \\ &\quad - \left. \int_0^{t_M} \int_0^t k(t-s) \int_0^1 G_1 \left(ry + (1-r)Y\right) dre(s)dsz(t)dt \right| \\ &\leq C\tilde{L} \int_0^{t_M} \int_0^t |k(t-s)|(e(s))^2 ds |z(t)| dt \\ &\leq C\tilde{L} \left(\int_0^{t_M} \left(\int_0^t |k(t-s)(e(s))^2 ds \right)^2 dt \right)^{1/2} \left(\int_0^{t_M} z^2(t)dt \right)^{1/2} \\ &\leq C\tilde{L} \int_0^{t_M} |k(t)|dt \left(\int_0^{t_M} e^4(t)dt \right)^{1/2} \left(\int_0^{t_M} z^2(t)dt \right)^{1/2} \\ &\leq C\tilde{L}\tilde{B} \left(\int_0^{t_M} e^2(t)dt \right)^{1/2} \left(\int_0^{t_M} z^2(t)dt \right)^{1/2} \cdot ||e(t)||_{[0,t_M]} \end{aligned} \tag{2.2.2.15}$$

We define $\wp$ to be the solution of

$$\begin{cases} -\wp' + a(t)\wp = \int_t^{t_M} k(s-t)A(s)\wp(s)ds, \ t_M > t > 0, \\ \wp(t_M) = e_M^-/|e_M^-|, \end{cases}$$

That is, $\wp = z/|e_M^-|$. So from (2.2.2.15)

$$\left|(D-\tilde{D})(e,z)\right| \leq C\tilde{L}\tilde{B}\,|e_M^-| \left(\int_0^{t_M} e^2 dt\right)^{1/2} \left(\int_0^{t_M} \wp^2 dt\right)^{1/2} ||e(t)||_{[0,t_M]}. \quad (2.2.2.16)$$

Next, we conclude that there exists a $X \in \mathcal{V}_N^{(m)}$ such that the first term of (2.2.2.13) satisfies

$$\left|\tilde{D}(e, z-X)\right| \leq \max_{n\leq M} \Re_m(Y, M; n). \quad (2.2.2.17)$$

Now we will now prove (2.2.2.17). Because of $y' + a(t)y - \mathcal{V}_G(Y)(t) = 0,\ [y]_n = 0$, (for all n) and $Y_0^- = y_0$, we find

$$\begin{aligned} \tilde{D}(e, z-X) &= -\sum_{n=1}^{M} \int_{I_n} (Y' + a(t)Y - \mathcal{V}_G(Y)(t))\,(z-X)dt \\ &- \sum_{n=1}^{M} [Y]_{n-1}(z-X)_{n-1}^{+}. \end{aligned}$$

Hence,

$$\begin{aligned} &\left|\tilde{D}(e, z-X)\right| \\ \leq\ &\left|\sum_{n=1}^{M} \int_{I_n} Y'(z-X)dt\right| + \left|\sum_{n=1}^{M} \int_{I_n} [a(t)Y - \mathcal{V}_G(Y)(t)](z-X)dt\right| \\ +\ &\sum_{n=1}^{M} \left|[Y]_{n-1}(z-X)_{n-1}^{+}\right| \\ =:\ &I + II + III. \end{aligned}$$

The remaining lines are then the same as in the proof of (2.1.2.6).

Thus, from (2.2.2.13), (2.2.2.15), (2.2.2.17), we derive

$$\begin{aligned} |e_M^-| \ &\leq \ \frac{\max_{n\leq M} \Re_m(Y, M; n)}{|e_M^-|} \\ &+ \ C\tilde{L}\tilde{B}||e||_{[0,t_M]} \left(\int_0^{t_M} e^2 dt\right)^{1/2} \cdot \left(\int_0^{t_M} \wp^2 dt\right)^{1/2}. \end{aligned} \quad (2.2.2.18)$$

Since G satisfies (2.2.2.8) and z is the solution of (2.2.2.10), following the proof of Lemma 2.1.2.1 leads to

$$S(M) = \int_0^{t_M} |z'|\, dt \leq C \left|e_M^-\right|, \tag{2.2.2.19}$$

where $C := C(t_M, \mathcal{A}, \mathcal{B}, L)$. Thus, combining (2.2.2.14), (2.2.2.19), and (2.2.2.19), we obtain

$$\begin{aligned} \left|e_M^-\right| &\leq \max_{n\leq M} C\left(|[Y]_{n-1}| + |a(t)Y - \mathcal{V}_G(Y)(t)|\, h_n\right) \\ &+ C\tilde{L}\tilde{B}||e||_{[0,t_M]} \left(\int_0^{t_M} e^2 dt\right)^{1/2} \left(\int_0^{t_M} \wp^2 dt\right)^{1/2}. \end{aligned} \tag{2.2.2.20}$$

To complete the proof of the theorem, we need the following lemma.

Lemma 2.2.2.1. *Let $\mathcal{A} := ||a||_I$, $\tilde{\mathcal{B}} := \int_0^{t_M} |k(t)|dt$, and suppose that G is Lipschitz continuous, i.e.,*

$$|G(u) - G(v)| \leq L|u - v|,$$

for all u, $v \in \Omega$. Then the error of DG(0) to (2.2.1.1) satisfies

$$||e||^2_{[0,t_M]} \leq C\left|e_{M-1}^-\right|^2 + C\left|[Y]_{M-1}\right|^2 + C\left(a(t)Y + \mathcal{V}_G(Y)(t)\right)^2,$$

where $C := C(t_M, L, \mathcal{A}, \tilde{\mathcal{B}})$ is independent of the mesh size h_n.

Proof. For $m = 0$, we have the following identity on I_n,

$$Y' + a(t)Y - \mathcal{V}_G(Y)(t) = a(t)Y - \mathcal{V}_G(Y)(t).$$

Subtracting this from (2.2.1.1) leads to

$$e' + a(t)e - \left(\int_0^t k(t-s)(G(y) - G(Y))ds\right) = -a(t)Y + \mathcal{V}_G(Y)(t). \tag{2.2.2.21}$$

Consequently, we have

$$e'e + a(t)e^2 - \left(\int_0^t k(t-s)(G(y) - G(Y))ds\right)e = -a(t)Ye + \mathcal{V}_G(Y)(t)e,$$

Integrating from t_{n-1} to t, we obtain

$$\int_{t_{n-1}}^{t}\frac{1}{2}\frac{d(e^2)}{dt}+\int_{t_{n-1}}^{t}a(t)e^2dt-\int_{t_{n-1}}^{t}\left(\int_0^t k(t-s)(G(y)-G(Y))ds\right)e\,dt$$
$$=\int_{t_{n-1}}^{t}(-a(t))Ye\,dt+\int_{t_{n-1}}^{t}\mathcal{V}_G(Y)(t)e\,dt,$$

and hence

$$\begin{aligned}
\frac{1}{2}(e^2) &\le \frac{1}{2}(e_{n-1}^{+})^2+\int_{I_n}|a(t)Ye|\,dt+\int_{I_n}|\mathcal{V}_G(Y)(t)\cdot e|\,dt\\
&+ \left(\int_{I_n}\left(\int_0^t k(t-s)(G(y)-G(Y))ds\right)^2dt\right)^{1/2}\\
&\cdot \left(\int_{I_n}e^2dt\right)^{1/2}+\int_{I_n}|a(t)|e^2dt\\
&\le \frac{1}{2}(e_{n-1}^{+})^2+\int_{I_n}|a(t)Ye|\,dt+\int_{I_n}|\mathcal{V}_G(Y)(t)\cdot e|\,dt\\
&+ \sqrt{2}\int_0^{t_n}|k(t)|dt\cdot\left(\int_0^{t_n}(G(y)-G(Y))^2dt\right)^{1/2}\\
&\cdot \left(\int_{I_n}e^2dt\right)^{1/2}+\int_{I_n}|a(t)|e^2dt\\
&\le \frac{1}{2}(e_{n-1}^{+})^2+\int_{I_n}|a(t)Ye|dt+\int_{I_n}|\mathcal{V}_G(Y)(t)\cdot e|\,dt\\
&+ \sqrt{2}T\tilde{B}L||e||^2_{[0,t_n]}h_n+\int_{I_n}|a(t)|e^2dt\\
&\le \left|e_{n-1}^{-}\right|^2+|[Y]_{n-1}|^2+\frac{h_n}{2}\left(a(t)Y+\mathcal{V}_G(Y)(t)\right)^2\\
&+ \frac{h_n}{2}||e||^2_{I_n}+\mathcal{A}h_n||e||^2_{[0,t_M]}+\sqrt{2}T\tilde{B}Lh_n||e||^2_{[0,t_M]}. \qquad (2.2.2.22)
\end{aligned}$$

For sufficiently small $h_n(\frac{1}{2}+\mathcal{A}+\sqrt{2}T\tilde{B}L)$ we derive that, from (2.2.2.22),

$$||e||^2_{[0,t_M]}\le C\left|e_{M-1}^{-}\right|^2+C\,|[Y]_{M-1}|^2+C\left(a(t)Y+\mathcal{V}_G(Y)(t)\right)^2. \qquad (2.2.2.23)$$

According to Theorem 2.2.2.1 and the procedure of the proof of Lemma 2.1.2.1, we know that, for sufficiently small h_n,

$$C\tilde{L}\tilde{B}\left(\int_0^{t_M}e^2dt\right)^{1/2}\cdot\left(\int_0^{t_M}\wp^2dt]\right)^{1/2}\le\frac{1}{2}. \qquad (2.2.2.24)$$

Combining (2.2.2.20), Lemma 2.2.2.1, and (2.2.2.24), we arrive at

$$\begin{aligned}||e||^2_{[0,t_M]} &\leq \max_{n\leq M} C\left(|[Y]_{n-1}| + h_n\, |a(t)Y - \mathcal{V}_G(Y)(t)|\right)^2 \\ &+ C([Y]_{M-1})^2 + C\left(a(t)Y + \mathcal{V}_G(Y)(t)\right)^2.\end{aligned}$$

2.2.3 Efficiency of the a posteriori error estimator

We shall use the a priori error estimates to deduce the efficiency of the a posteriori estimator described in Theorem 2.2.2.1.

Theorem 2.2.3.1. *Under the assumptions of Theorem 2.2.2.2 and Theorem 2.2.1.1, we have*

$$\left|y(t_M) - Y_M^-\right| \leq C||h_n^{m+1}R(Y)||_{[0,t_M]} \leq Ch^{m+1}(||y'(t)||_{[0,t_M]} + ||y^{(m+1)}(t)||_{[0,t_M]})$$

where $C := C(t_M, L, \mathcal{A}, \mathcal{B})$ is independent of the mesh size h.

Proof. We only need to bound the term $||h_n^{m+1}R(Y)||_{[0,t_M]}$ in Theorem 2.2.2.1:

$$\begin{aligned}||h_n^{m+1}R(Y)||_{[0,t_M]} &\leq ||h_n^m[Y]_{n-1}||_{[0,t_M]} + ||h_n^{m+1}(a(t)Y - \mathcal{V}_G(Y)(t))||_{[0,t_M]} \\ &= ||h_n^m(Y_{n-1}^+ - y + y - Y_{n-1}^-)||_{[0,t_M]} + ||h_n^{m+1}(-y'(t) - a(t)y(t) \\ &+ \mathcal{V}_G(y)(t) + a(t)Y - \mathcal{V}_G(Y)(t))||_{[0,t_M]} \\ &\leq 2||h_n^m e(t)||_{[0,t_M]} + ||h_n^{m+1}y'(t)||_{[0,t_M]} + ||h_n^{m+1}[-a(t)e(t) \\ &+ \int_0^t k(t-s)(G(y(s)) - G(Y(s)))ds]||_{[0,t_M]} \\ &\leq 2h_n^m||e(t)||_{[0,t_M]} + h^{m+1}||y'(t)||_{[0,t_M]} \\ &+ (\mathcal{A}h^{m+1} + Lh^{m+1}t_M)\mathcal{B}||e(t)||_{[0,t_M]}.\end{aligned}$$

Then we combine this with the result of Theorem 2.2.1.1 and complete the proof.

2.3 Numerical examples

In this section, we compare the accuracy and stability of DG(0), CG(1) and the collocation method using piecewise linear polynomial approximation (denoted by CC(1)) by means of numerical examples. The effect of quadrature on the total error is considered, too.

2.3.1 Example for the case of constant coefficient

Example 2.3.1.1. *We consider the linear scalar Volterra integro-differential equation*

$$y' + ay = \int_0^t \exp(-(t-s))y(s)ds, \quad t \in I = [0,1], \quad y(0) = 1, \tag{2.3.1.1}$$

where a is a constant.

The exact solution of (2.3.1.1) is

$$\begin{aligned} y(t) &= \exp(-\frac{a+1}{2}t)\cosh\left(\sqrt{1-a+\frac{(a+1)^2}{4}}t\right) \\ &+ \frac{1-a}{\sqrt{a^2-2a+5}}\exp(-\frac{a+1}{2}t)\sinh\left(\sqrt{1-a+\frac{(a+1)^2}{4}}t\right). \end{aligned} \tag{2.3.1.2}$$

We take uniform meshes: $\{t_i : t_i = ih, i = 0, 1, \cdots, n\}$, where h is the mesh size, and the initial value is $Y_0^- = y(0) = 1$.

DG(0):

$$\begin{aligned} & (ah - h + 2 - \exp(-h))Y_n^- = Y_{n-1}^- \\ + & \sum_{i=1}^{n-1} Y_i^-[-\exp(-(t_n - t_i)) + \exp(-(t_n - t_{i-1})) \\ + & \exp(-(t_{n-1} - t_i)) - \exp(-(t_{n-1} - t_{i-1}))]. \end{aligned} \tag{2.3.1.3}$$

CG(1):

$$Y_n - Y_{n-1} + \int_{I_n} a(t)Y(t)dt = \int_{I_n}\int_0^t k(t-s)Y(s)dsdt, \qquad (2.3.1.4)$$

Taking $Y|_{I_n} = \frac{t-t_n}{-k}Y_{n-1} + \frac{t-t_{n-1}}{k}Y_n$ in (2.3.1.4), we obtain the exact CG(1) and CG(1) with quadrature scheme for Example 2.3.1.1 as follows.

Exact CG(1):

$$\begin{aligned}
& (2 - h/2 + ah/2 + \frac{1}{h}\exp(-h) - 1/h)Y_n \\
= \;& (1 - ah/2 + h/2 + \exp(-h) + \frac{1}{h}\exp(-h) - 1/h)Y_{n-1} \\
+ \;& \sum_{i=1}^{n-1} Y_{i-1}[\exp(-(t_n - t_{i-1})) - \frac{1}{h}\exp(-(t_n - t_i)) + \frac{1}{h}\exp(-(t_n - t_{i-1})) \\
- \;& \exp(-(t_{n-1} - t_{i-1})) + \frac{1}{h}\exp(-(t_{n-1} - t_i)) - \frac{1}{h}\exp(-(t_{n-1} - t_{i-1}))] \\
+ \;& \sum_{i=1}^{n-1} Y_i[\exp(-(t_n - t_i)) + \frac{1}{h}\exp(-(t_n - t_i)) - \frac{1}{h}\exp(-(t_n - t_{i-1})) \\
+ \;& \exp(-(t_{n-1} - t_i)) - \frac{1}{h}\exp(-(t_{n-1} - t_i)) \\
+ \;& \frac{1}{h}\exp(-(t_{n-1} - t_{i-1}))]. \qquad (2.3.1.5)
\end{aligned}$$

From [23], we know that the collocation method using piecewise linear polynomial, i.e., CC(1) to (2.1.1.1), has the form

$$\begin{aligned}
& (1 + ha(t_{n+1}) - h^2\int_0^1 k(t_{n+1} - (t_n + sh))sds)Y_{n+1} \\
= \;& (1 + h^2\int_0^1 k(t_{n+1} - (t_n + sh))(1-s)ds)Y_n + hF_n(t_{n+1}),
\end{aligned}$$

where $F_n(t_{n+1}) = \sum_{i=0}^{n-1} h\int_0^1 k(t_{n+1} - (t_i + sh))[(1-s)Y_i + sY_{i+1}]ds$.

Exact CC(1):

$$\begin{aligned} & (2+ah-h-\exp(-h))Y_{n+1} \\ = & \ (2-h\exp(-h)-\exp(-h))Y_n \\ + & \ h\sum_{i=0}^{n-1}[\exp(-(t_{n+1}-t_i))(\exp(h)-\exp(h)/h+1/h)Y_{i+1} \\ + & \ [-\exp(-(t_{n+1}-t_i))+\frac{1}{h}\exp(-(t_{n+1}-t_i))(\exp(h)-1)]Y_i]. \quad (2.3.1.6) \end{aligned}$$

We take the end-point rule for the inner product and the trapezoidal rule for the memory term, and call it "Quadrature Scheme 1".

Quadrature Scheme 1 for DG(0):

$$\begin{aligned} & [ah+1-h^2/2-h^2\exp(-h)/2]Y_n^- \\ = & \ Y_{n-1}^- + \sum_{i=1}^{n-1} Y_i^- \frac{h^2}{2}[\exp(-(t_n-t_i))+\exp(-(t_n-t_{i-1}))]. \quad (2.3.1.7) \end{aligned}$$

Quadrature Scheme 1 for CG(1):

$$\begin{aligned} (1+ah-h^2/2)Y_n &= (1+\frac{h^2}{2}\exp(-h))Y_{n-1} + \sum_{i=1}^{n-1} Y_{i-1}\frac{h^2}{2}\exp(-(t_n-t_{i-1})) \\ &+ \sum_{i=1}^{n-1} Y_i \frac{h^2}{2}\exp(-(t_n-t_i)). \quad (2.3.1.8) \end{aligned}$$

Quadrature Scheme 1 for CC(1):

$$\begin{aligned} (1+ah-h^2/2)Y_{n+1} &= (1+h^2\exp(-h)/2)Y_n + h\sum_{i=0}^{n-1}\frac{h}{2}\exp(-(t_{n+1}-t_i))Y_i \\ &+ h\sum_{i=0}^{n-1}\frac{h}{2}\exp(-(t_{n+1}-t_i-h))Y_{i+1}. \quad (2.3.1.9) \end{aligned}$$

The "Quadrature Scheme 2" is defined by taking the end point rule for the inner product and the mid-point rule for the memory term.

Quadrature Scheme 2 for DG(0):

$$[ah+1-h^2\exp(-h/2)]Y_n^- = Y_{n-1}^- + \sum_{i=1}^{n-1} Y_i^- h^2\exp(-(t_n-\frac{t_i+t_{i-1}}{2})). \quad (2.3.1.10)$$

Quadrature Scheme 2 for CG(1):

$$
\begin{aligned}
& [1 + ah - h^2 \exp(-h/2)/2]Y_n \\
= \;& [h^2 \exp(-h/2)/2 + 1]Y_{n-1} + \frac{h^2}{2}\sum_{i=1}^{n-1} Y_{i-1} \exp(-(t_n - \frac{t_i + t_{i-1}}{2})) \\
+ \;& \frac{h^2}{2}\sum_{i=1}^{n-1} Y_i \exp(-(t_n - \frac{t_i + t_{i-1}}{2})).
\end{aligned}
\tag{2.3.1.11}
$$

Quadrature Scheme 2 for CC(1):

$$
\begin{aligned}
& (1 + ah - h^2 \exp(-h/2)/2)Y_{n+1} \\
= \;& (1 + \frac{h^2}{2}\exp(-h/2))Y_n + h^2/2\sum_{i=0}^{n-1} Y_i \exp(-(t_{n+1} - t_i - h/2)) \\
+ \;& h^2/2\sum_{i=0}^{n-1} Y_{i+1} \exp(-(t_{n+1} - t_i - h/2)).
\end{aligned}
\tag{2.3.1.12}
$$

First of all, we compare the accuracy of DG(0), CG(1) and CC(1) through Example 2.3.1.1 with $a = 7$, using the same method as [44]. Other values of a lead to similar results. We assume that the maximum norm of the error in I is proportional to h^p, that is, $|\text{error}| \approx Ch^p$ with the constant of C independent of meshsize h. To determine the order p experimentally, we take logarithms:

$$\log(|\text{error}|) \approx \log(C) + p\log(h),$$

By doing so we can determine p as the slope of a line that passes through the points $(\log(h), \log(|\text{error}|))$. We plot the logarithms of the errors versus the logarithms of the corresponding time steps for exact scheme in Figure 2.1. The slopes of the lines are 0.9837 for DG(0), 0.9836 for CC(1) and 2.0004 for CG(1). Correspondingly in Figure 2.2 for Quadrature Scheme 1, the line slope for DG(0) is 0.9861, for CG(1), 0.9414, for CC(1), 0.9836, which reveals that they have the same order of convergence. However, the distance between the lines indicates that the constant

C in the representation of the error, $|\text{error}| \approx Ch^p$, for Quadrature Scheme 1 is different from the exact DG(0). The same explanation holds for Quadrature Scheme 2 in Figure 2.3. Figure 2.4 illustrates that the stability of DG(0) and CC(1) is better than that of CG(1).

2.3.2 Example of a VIDE with time-dependent coefficient

Example 2.3.2.1. *We consider the linear scalar Volterra integro-differential equation*

$$y' + a(t)y = \int_0^t \exp\left(-(t-s)\right)y(s)ds, \quad t \in I = (0,1], \quad y(0) = 1, \tag{2.3.2.1}$$

where $a(t) = r + \frac{1}{1-r} - \frac{1}{1-r}\exp\left((r-1)t\right)$ *and* $r \neq 1$ *is a positive constant,*

The exact solution of (2.3.2.1) is

$$y(t) = \exp(-rt). \tag{2.3.2.2}$$

We take uniform meshes: $\{t_i : \ t_i = ih, \ i = 0, 1, \cdots, n\}$ and choose the initial value $Y_0^- = y(0) = 1$.

We easily formulate the schemes of DG(0), CG(1) and CC(1) for Example 2.3.2.1, as follows.

Exact DG(0):

$$\begin{aligned} & (h(r + \frac{1}{1-r}) + \frac{1}{(r-1)^2}[\exp\left((r-1)t_n\right) \\ - \ & \exp\left((r-1)t_{n-1}\right)] - h + 2 - \exp\left(-h\right))Y_n^- \\ = \ & Y_{n-1}^- + \sum_{i=1}^{n-1} Y_i^-[-\exp\left(-(t_n - t_i)\right) + \exp\left(-(t_n - t_{i-1})\right) \\ + \ & \exp\left(-(t_{n-1} - t_i)\right) - \exp\left(-(t_{n-1} - t_{i-1})\right)]. \end{aligned} \tag{2.3.2.3}$$

Exact CG(1):

$$
\begin{aligned}
& (2 - h/2 + (r + \frac{1}{1-r})\frac{h}{2} + \frac{1}{(r-1)^2}\exp((r-1)t_n) - \frac{1}{h}\frac{1}{(r-1)^3}\exp((r-1)t_n) \\
+ & \frac{1}{h}\frac{1}{(r-1)^3}\exp((r-1)t_{n-1}) + \frac{1}{h}\exp(-h) - 1/h)Y_n \\
= & [1 - [(r + \frac{1}{1-r})\frac{h}{2} - \frac{1}{(r-1)^2}\exp((r-1)t_{n-1}) + \frac{1}{h}\frac{1}{(r-1)^3}\exp((r-1)t_n) \\
- & \frac{1}{h}\frac{1}{(r-1)^3}\exp((r-1)t_{n-1})] + h/2 + \exp(-h) + \frac{1}{h}\exp(-h) - 1/h]Y_{n-1} \\
+ & \sum_{i=1}^{n-1} Y_{i-1}[\exp(-(t_n - t_{i-1})) - \frac{1}{h}\exp(-(t_n - t_i)) + \frac{1}{h}\exp(-(t_n - t_{i-1})) \\
- & \exp(-(t_{n-1} - t_{i-1})) + \frac{1}{h}\exp(-(t_{n-1} - t_i)) - \frac{1}{h}\exp(-(t_{n-1} - t_{i-1}))] \\
+ & \sum_{i=1}^{n-1} Y_i[-\exp(-(t_n - t_i)) + \frac{1}{h}\exp(-(t_n - t_i)) - \frac{1}{h}\exp(-(t_n - t_{i-1})) \\
+ & \exp(-(t_{n-1} - t_i)) - \frac{1}{h}\exp(-(t_{n-1} - t_i)) \\
+ & \frac{1}{h}\exp(-(t_{n-1} - t_{i-1}))].
\end{aligned}
\tag{2.3.2.4}
$$

Exact CC(1):

$$
\begin{aligned}
& (2 + h(r + \frac{1}{1-r} - \frac{1}{1-r}\exp((r-1)t_{n+1})) - h - \exp(-h))Y_{n+1} \\
= & (2 - h\exp(-h) - \exp(-h))Y_n + h\sum_{i=0}^{n-1}\{\exp(-(t_{n+1} - t_i)) \\
\cdot & (\exp(h) - \exp(h)/h + 1/h)Y_{i+1} + [-\exp(-(t_{n+1} - t_i)) \\
+ & \frac{1}{h}\exp(-(t_{n+1} - t_i))(\exp(h) - 1)]Y_i\}.
\end{aligned}
\tag{2.3.2.5}
$$

Now we consider the "Quadrature Scheme 2" for DG(0), CG(1) and CC(1).

Quadrature Scheme 2 for DG(0):

$$
\begin{aligned}
& [h(r + \frac{1}{1-r} - \frac{1}{1-r}\exp((r-1)t_n)) + 1 - h^2\exp(-h/2)]Y_n^- \\
= & Y_{n-1}^- + \sum_{i=1}^{n-1} Y_i^- h^2 \exp(-(t_n - \frac{t_i + t_{i-1}}{2})).
\end{aligned}
\tag{2.3.2.6}
$$

Quadrature Scheme 2 for CG(1):

$$
\begin{aligned}
& [1 + h(r + \frac{1}{1-r} - \frac{1}{1-r}\exp((r-1)t_n)) - h^2 \exp(-h/2)/2]Y_n \\
= \; & [h^2 \exp(-h/2)/2 + 1]Y_{n-1} + \frac{h^2}{2}\sum_{i=1}^{n-1} Y_{i-1}\exp(-(t_n - \frac{t_i + t_{i-1}}{2})) \\
+ \; & \frac{h^2}{2}\sum_{i=1}^{n-1} Y_i \exp(-(t_n - \frac{t_i + t_{i-1}}{2})).
\end{aligned} \tag{2.3.2.7}
$$

Quadrature Scheme 2 for CC(1):

$$
\begin{aligned}
& (1 + h(r + \frac{1}{1-r} - \frac{1}{1-r}\exp((r-1)t_{n+1})) - h^2 \exp(-h/2)/2)Y_{n+1} \\
= \; & (1 + \frac{h^2}{2}\exp(-h/2))Y_n + h^2/2\sum_{i=0}^{n-1} Y_i \exp(-(t_{n+1} - t_i - h/2)) \\
+ \; & h^2/2\sum_{i=0}^{n-1} Y_{i+1}\exp(-(t_{n+1} - t_i - h/2)).
\end{aligned} \tag{2.3.2.8}
$$

When we take $r = 6$, we obtain the numerical results shown in Figure 2.5 and Figure 2.6, whose explanation is the same as for Example 2.3.1.1. Other values of r lead to similar results.

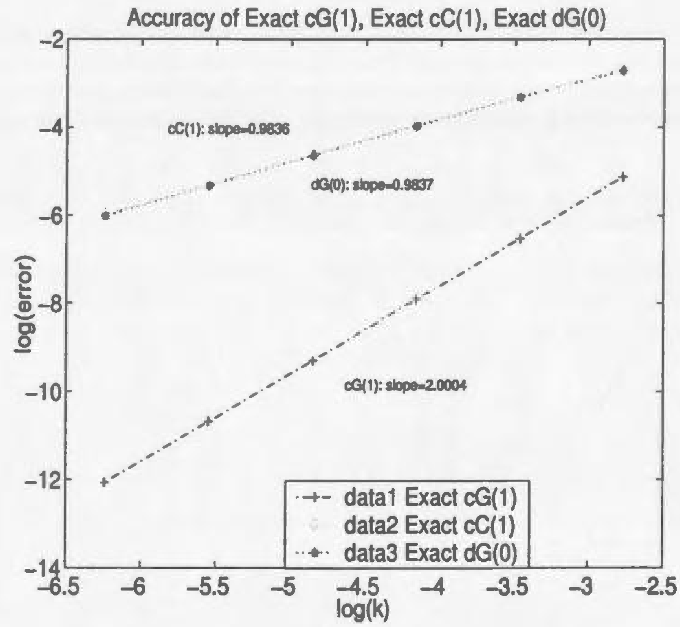

Figure 2.1: The methods for Example 2.3.1.1

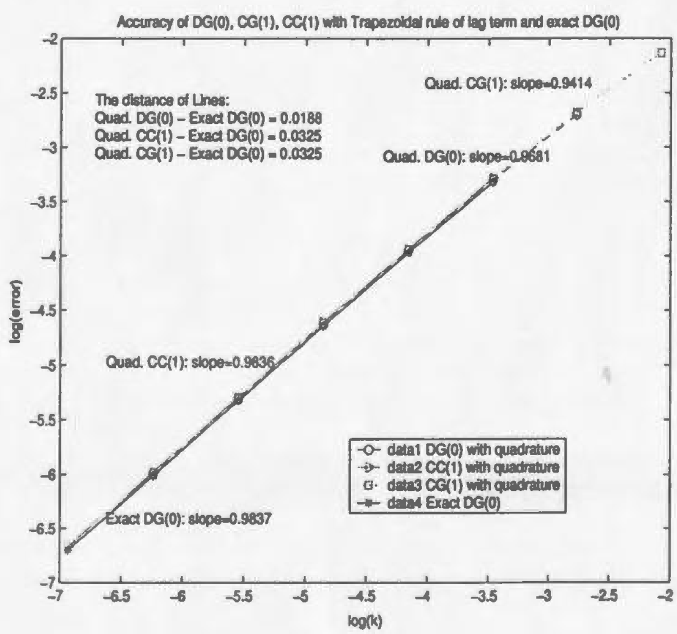

Figure 2.2: The methods with Trapezoidal rule of lag term for Example 2.3.1.1

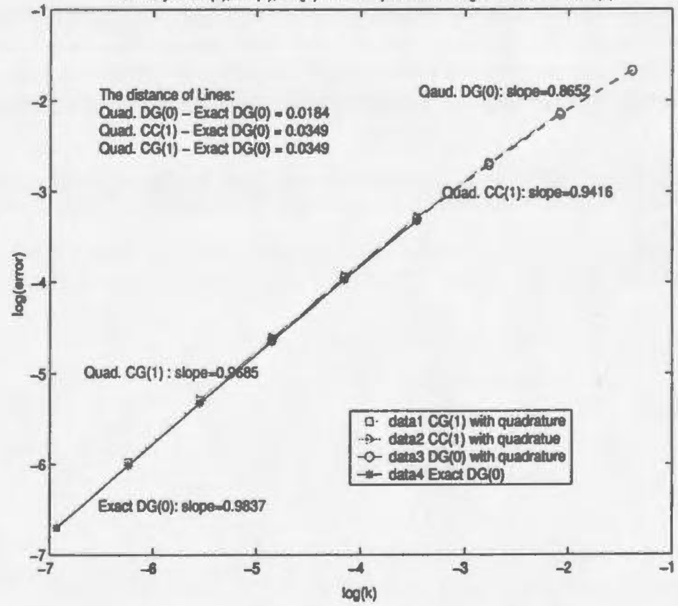

Figure 2.3: The methods with midpoint rule of lag term for Example 2.3.1.1

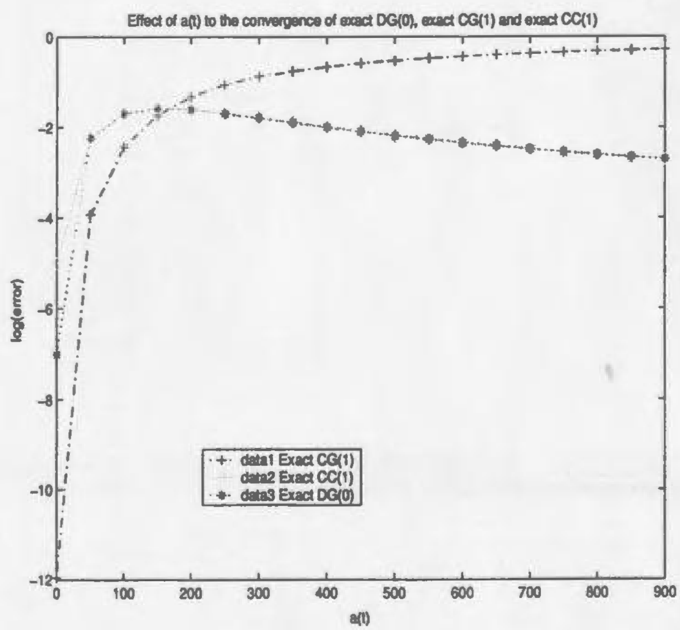

Figure 2.4: Effect of stability on the convergence of methods for Example 2.3.1.1

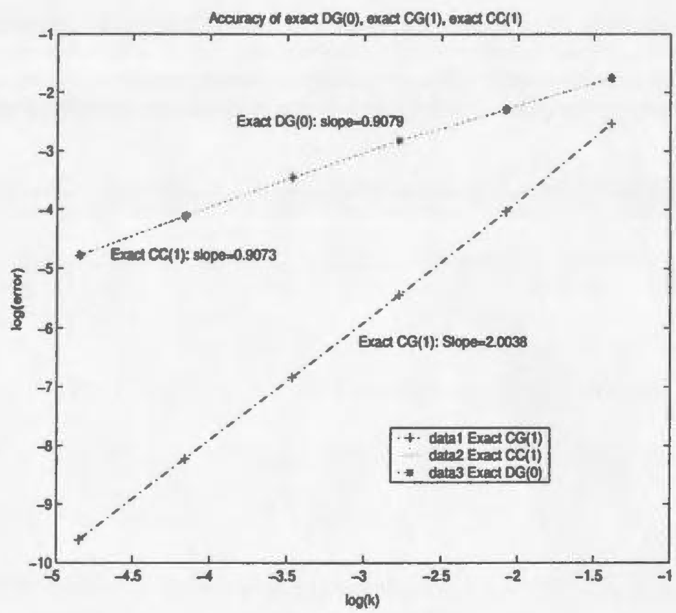

Figure 2.5: The methods for Example 2.3.2.1

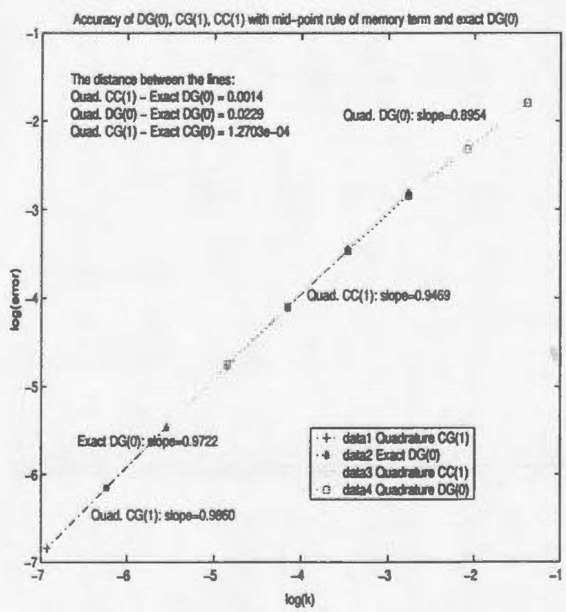

Figure 2.6: The methods with midpoint rule of lag term for Example 2.3.2.1

2.4 The discontinuous Galerkin method for nonstandard VIDEs

2.4.1 Preliminaries

In this section, we study the nonstandard Volterra integro-differential equation,

$$\begin{cases} y' + a(t)y = \mathcal{V}_G^N(y)(t), & t \in I := [0,T], \\ y(0) = y_0, \end{cases} \tag{2.4.1.1}$$

where $\mathcal{V}_G^N(y)(t) := \int_0^t k(t-s)G(y(t), y(s))ds$ and $a,\ k \in C(I)$. Assume that the (Lipschitz continuous) function $G : \Omega \times \Omega \to \mathbb{R}$ ($\Omega \subset \mathbb{R}$) is such that (2.4.1.1) possesses a unique solution $y \in C^1(I)$ for all $y_0 \in \Omega$.

We give the a posteriori error estimates of DG(m) to (2.4.1.1). As in the above sections, we write the DG scheme as

$$B_{DG}(Y, X) := F_{DG}(X), \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.4.1.2}$$

where

$$\begin{aligned} B_{DG}(Y,X) &:= \sum_{n=1}^{N} \int_{I_n} \{Y'(t)X(t) + a(t)Y(t)X(t) & (2.4.1.3) \\ &- \mathcal{V}_G^N(Y)(t)X(t)\}dt + \sum_{n=1}^{M-1} [Y]_n X_n^+ + Y_0^+ X_0^+, \\ F_{DG}(X) &:= y_0 X_0^+. & (2.4.1.4) \end{aligned}$$

To show that (2.4.1.2) has a unique solution $Y|_{I_n} \in \mathcal{P}^{(m)}(I_n)$ we define that for $\tilde{Y} \in \mathcal{P}^{(m)}(I_n)$, $Y = \tilde{T}\tilde{Y} \in \mathcal{P}^{(m)}(I_n)$ as the solution of

$$\begin{aligned} &\int_{I_n} \{Y'(t)X(t) + a(t)Y(t)X(t) - (\int_0^t k(t-s)G(\tilde{Y}(t), Y(s))ds)X(t)\}dt \\ +\ & Y_{n-1}^+ X_{n-1}^+ = Y_{n-1}^- X_{n-1}^+, \end{aligned} \tag{2.4.1.5}$$

for all $X(t) \in \mathcal{P}^{(m)}(I_n)$. If the operator $\tilde{T}$ is a contraction on $\mathcal{P}^{(m)}(I_n)$ for all sufficiently small h_n, the assertion follows from Banach's fixed point theorem.

2.4.2 A posteriori error estimates for nonstandard VIDEs

Theorem 2.4.2.1. *Let $\mathcal{A} := ||a||_I$, $\mathcal{B} := ||k||_I$, and let $m = 0,\ 1$. Assume that G satisfies*

$$|\nabla G(u,v)| \leq L, \quad \forall u, v \in \Omega.$$

Then the error of the DG(m) approximation to (2.4.1.1) satisfies

$$|y(t_M) - Y_M^-| \leq C||h_n^{m+1} R(Y)||_{[0,t_M]},$$

where $C := C(t_M, \mathcal{A}, \mathcal{B}, L)$ and $R(Y) := \frac{|[Y]_{n-1}|}{h_n} + \left|a(t)Y - \mathcal{V}_G^N(Y)(t)\right|$ $(t \in I_n)$.

Proof. From (2.4.1.2), we know that

$$B_{DG}(Y,X) - B_{DG}(y,X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}.$$

We write this as

$$\tilde{D}(e,X) = 0, \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.4.2.6}$$

where

$$\begin{aligned} \tilde{D}(W,X) \quad := \quad & \sum_{n=1}^{M} \int_{I_n} \{W'X + a(t)WX - \int_0^t k(t-s) \qquad (2.4.2.7) \\ \cdot \quad & (\int_0^1 \nabla G(ry(t) + (1-r)Y(t), ry(s) + (1-r)Y(s))dr \\ \cdot \quad & (W(t), W(s)))ds \cdot X(t)\}dt + \sum_{n=1}^{M-1} [W]_n X_n^+ + W_0^+ X_0^+. \end{aligned}$$

We know that for all piecewise continuous w

$$\tilde{D}(w,z) = w_M^- e_M^-, \tag{2.4.2.8}$$

where z is the solution of the continuous linearized dual problem of (2.4.1.1), defined as

$$\begin{cases} -z' + [a(t) + \int_0^t k(t-s) \int_0^1 G_1(ry(t) + (1-r)Y(t), ry(s) \\ +(1-r)Y(s))drds]z = \int_t^{t_M} k(t-s)\check{A}(s)z(s)ds, \quad t_M > t > 0, \\ z(t_M) = e_M^-. \end{cases} \tag{2.4.2.9}$$

Here, $\check{A}(s) := \int_0^1 G_2(ry(t) + (1-r)Y(t), ry(s) + (1-r)Y(s))dr$, $\nabla G := (G_1, G_2)$.

Selecting $w = e := y - Y$ in (2.4.2.8), we get

$$\begin{aligned} [e_M^-]^2 &= \tilde{D}(e, z) = \tilde{D}(e, z - X) \\ &= B_{DG}(y, z - X) - B_{DG}(Y, z - X) \\ &= -\sum_{n=1}^{M} \int_{I_n} Y'(t)(z - X)dt - \sum_{n=1}^{M} \{a(t)Y(t)(z - X) \\ &- \mathcal{V}_G^N(Y)(t)(z - X)\}dt - \sum_{n=1}^{M} [Y]_{n-1}(z - X)_{n-1}^+, \\ &=: I + II + III, \end{aligned} \tag{2.4.2.10}$$

From the definition of (2.1.2.7), we easily derive

$$I = 0, \ |III| \leq S(M) \max_{n \leq M} |[Y]_{n-1}|, \tag{2.4.2.11}$$

$$\begin{aligned} |II| &:= |\sum_{n=1}^{M} [a(t)Y(t) - \mathcal{V}_G^N(Y)(t)](z - X)dt|, \\ &\leq \sum_{n=1}^{M} h_n ||a(t)Y(t) - \mathcal{V}_G^N(Y)(t)||_{\bar{I}_n} ||z - X||_{\bar{I}_n}, \\ &\leq \max_{n \leq M} h_n ||a(t)Y - \mathcal{V}_G^N(Y)(t)||_{\bar{I}_n} \cdot \sum_{n=1}^{M} h_n^{m+1} \int_{I_n} |z'|dt, \\ &= S(M) \max_{n \leq M} h_n ||a(t)Y - \mathcal{V}_G^N(Y)(t)||_{\bar{I}_n}, \end{aligned} \tag{2.4.2.12}$$

where $S(M) := \int_0^{t_M} |z^{(m+1)}|dt$. We need the stability of (2.4.2.9),

$$S(M) = \int_0^{t_M} |z^{(m+1)}|dt \leq C|e_M^-|, \tag{2.4.2.13}$$

where $C := C(t_M, \mathcal{A}, \mathcal{B}, L)$. Because (2.4.2.9) is the linearized dual problem, the proof of (2.4.2.13) is very similar to that of Lemma 2.1.2.1.

Combining the estimates (2.4.2.11), (2.4.2.12), and (2.4.2.13) with (2.4.2.10), we complete the proof of Theorem 2.4.2.1.

2.4.3 Efficiency of the a posteriori error estimator

We shall first derive the a priori error estimate of (2.4.1.1). This is then used to prove the efficiency of the a posteriori error estimator in Theorem 2.4.2.1.

Theorem 2.4.3.1. *Define* $\mathcal{A} := ||a||_I$, $\mathcal{B}^* := (\int_{I_n} |k(t)|dt)^2$, *and assume* G *satisfies*

$$|\nabla G(u, v)| \leq L, \quad \forall u, \ v \in \Omega.$$

Then the error of the DG(m) approximation to (2.4.1.1) satisfies

$$||e(t)||_{[0,t_M]} \leq C \max_{n \leq M} h_n^{m+1} ||y^{(m+1)}||_{\bar{I}_n},$$

with $m = 0, \ 1,$ *and* $C := C(t_M, L, \mathcal{A}, \mathcal{B}^*)$.

Proof. The proof is very similar to that of Theorem 2.2.1.1 and is thus left to the reader.

Theorem 2.4.3.2. *Under the assumptions of Theorem 2.4.2.1 and Theorem 2.4.3.1, we have*

$$|y(t_M) - Y_M^-| \leq C||h_n^{m+1} R(Y)||_{[0,t_M]} \leq Ch^{m+1}(||y'(t)||_{[0,t_M]} + ||y^{(m+1)}(t)||_{[0,t_M]}),$$

where $C := C(t_M, L, \mathcal{A}, \mathcal{B})$ *is independent of the mesh size* h.

Proof. We only need to bound the term $||h_n^{m+1}R(Y)||_{[0,t_M]}$ in Theorem 2.4.2.1:

$$\begin{aligned}
||h_n^{m+1}R(Y)||_{[0,t_M]} &\leq ||h_n^m[Y]_{n-1}||_{[0,t_M]} + ||h_n^{m+1}[a(t)Y - \mathcal{V}_G^N(Y)(t)]||_{[0,t_M]} \\
&= [h_n^m|Y_{n-1}^+ - y + y - Y_{n-1}^-|]_{[0,t_M]} + ||h_n^{m+1}(-y'(t) - a(t)y(t) \\
&+ \mathcal{V}_G^N(y)(t) + a(t)Y - \mathcal{V}_G^N(Y)(t))||_{[0,t_M]} \\
&\leq 2||h_n^m e(t)||_{[0,t_M]} + ||h_n^{m+1}y'(t)||_{[0,t_M]} + ||h_n^{m+1}[-a(t)e(t) \\
&+ \int_0^t k(t-s)(G(y(t),y(s)) - G(Y(t),Y(s)))ds]||_{[0,t_M]} \\
&\leq 2h^m||e(t)||_{[0,t_M]} + h^{m+1}|y'(t)| \\
&+ \mathcal{A}h^{m+1}||e(t)||_{[0,t_M]} + \sqrt{2}Lh^{m+1}t_M\mathcal{B}||e(t)||_{[0,t_M]}.
\end{aligned}$$

The proof is completed by recalling Theorem 2.4.3.1.

2.5 The discretized discontinuous Galerkin method for VIDEs

In this section we consider the discontinuous Galerkin methods with quadrature for the memory term and for the inner product for linear Volterra integro-differential equations. The readers are suggested to compare this section with Brunner [23].

2.5.1 The comparison with collocation method for VIDEs

We recall the DG time-stepping scheme for (2.1.1.1): For $n = 1, \cdots, N$, find $Y|_{I_n} \in \mathcal{P}^{(m)}(I_n)$, such that

$$\int_{I_n} (Y'(t) + a(t)Y(t) - \mathcal{V}(Y)(t))X dt + Y_{n-1}^+ X_{n-1}^+ = Y_{n-1}^- X_{n-1}^+, \qquad (2.5.1.1)$$

for all $X \in \mathcal{P}^{(m)}(I_n)$. Here we set $Y_0^- = y_0$. Suppose now that the integrals in (2.5.1.1) are approximated by interpolatory $(m+1)$-point quadrature formulas with abscissas $t_{n,j} := t_n + c_j h_n$ $(0 =: c_0 < c_1 < \ldots < c_m \leq 1)$ and weights w_j $(j =$

$0, 1, \ldots, m$). We denote the resulting *discretized* DG(m) solution in $\mathcal{V}_N^{(m)}$ by $\tilde{Y}$. The semi-discretized version of (2.5.1.1) is then given by

$$h_n \sum_{j=0}^{m} w_j[\tilde{Y}'(t_{n,j}) + a(t_{n,j})\tilde{Y}(t_{n,j}) - Z(t_{n,j})]X(t_{n,j}) + \tilde{Y}(t_n^+)X(t_n^+) - \tilde{Y}(t_n^-)X(t_n^-) = 0, \tag{2.5.1.2}$$

for all $X \in \mathcal{P}^{(m)}(I_n)$, where

$$\begin{aligned} Z(t_{n,j}) &:= \mathcal{V}(\tilde{Y})(t_{n,j}) \\ &= \int_0^{t_{n,j}} k(t_{n,j} - s)\tilde{Y}(s)ds \\ &:= F_n + \int_{t_n}^{t_{n,j}} k(t_{n,j} - s)\tilde{Y}(s)ds. \end{aligned}$$

We denote the discretized version of $Z(t_{n,j})$ by

$$\tilde{Z}(t_{n,j}) := \tilde{F}_n + h_n \sum_{\ell=0}^{j} w_{n,\ell} k(t_{n,j} - t_{n,l})\tilde{Y}(t_{n,l}). \tag{2.5.1.3}$$

The fully discretized version of (2.5.1.1) is then defined by

$$h_n \sum_{j=0}^{m} w_j[\tilde{Y}'(t_{n,j}) + a(t_{n,j})\tilde{Y}(t_{n,j}) - \tilde{Z}(t_{n,j})]X(t_{n,j}) + \tilde{Y}(t_n^+)X(t_n^+) - \tilde{Y}(t_n^-)X(t_n^-) = 0, \tag{2.5.1.4}$$

for all $X \in \mathcal{P}^{(m)}(I_n)$.

The above fully discretized DG method (2.5.1.4) is called:

(i) *an extended fully discretized* DG method if the lag term formula for F_n is given by

$$\tilde{F}_n := h_n \sum_{\ell=0}^{n-1} \sum_{j=1}^{m} w_{n,j} k(t_n - t_\ell + c_j h)\tilde{Y}_{\ell,j}, \quad n = 1, \ldots, N-1; \tag{2.5.1.5}$$

(ii) *a mixed fully discretized* DG method if the lag term formula is defined by

$$\tilde{F}_n := h_n \sum_{\ell=0}^{n-1} w_{n,\ell} k(t_n - t_\ell)\tilde{Y}_\ell, \quad n = 1, \ldots, N-1. \tag{2.5.1.6}$$

Let

$$\begin{aligned}
\tilde{Y}_n &:= \tilde{Y}(t_n^-), \\
\tilde{Y}_{n,0} &:= \tilde{Y}(t_n^+) \ (= \tilde{Y}(t_{n,0}^+)), \\
\tilde{Y}_{n,j} &:= \tilde{Y}(t_{n,j}) \ (j = 1, \dots, m), \\
\tilde{Z}_{n,j} &:= \tilde{Z}(t_{n,j}) \ (j = 1, \dots, m).
\end{aligned}$$

and let $L_j(v)$ be the jth Lagrange canonical polynomial (of degree $m-1$) corresponding to the points $\{c_i : i = 1, \dots, m\}$. Moreover, denote by $\{X_j : j = 0, 1, \dots, m\}$ a (canonical) basis for $\mathcal{P}^{(m)}(I_n)$ so that

$$X_i(t_n + c_j h_n) = \delta_{i,j} \quad (i,\ j = 0, 1, \dots, m).$$

Since the restriction of $\tilde{Y}'$ to I_n is a polynomial of degree $m-1$ we may write

$$\tilde{Y}'(t_n + v h_n) = \sum_{j=1}^{m} L_j(v) \tilde{Y}'(t_{n,j}), \quad v \in (0, 1],$$

and hence

$$\tilde{Y}(t_n + v h_n) = \tilde{Y}(t_n^+) + h_n \int_0^v \tilde{Y}'(t_n + s h_n) ds, \quad v \in (0, 1]. \tag{2.5.1.7}$$

On the other hand, (2.5.1.4) with $X = X_0$ yields

$$h_n w_0 [\tilde{Y}'(t_{n,0}) + a(t_{n,0}) \tilde{Y}(t_{n,0}) - \tilde{Z}(t_{n,0})] + \tilde{Y}(t_n^+) - \tilde{Y}(t_n^-) = 0,$$

implying that

$$\tilde{Y}(t_n^+) = \tilde{Y}_n + h_n w_0 [\tilde{Z}(t_{n,0}) - a(t_{n,0}) \tilde{Y}(t_n^+) - \sum_{j=1}^{m} L_j(c_0) \tilde{Y}'(t_{n,j})]. \tag{2.5.1.8}$$

For $X = X_i$ $(i = 1, \dots, m)$, with $X_i(t_{n,j}) = \delta_{i,j}$, we obtain from (2.5.1.4) the equations

$$w_i [\tilde{Y}'(t_{n,i}) + a(t_{n,i}) \tilde{Y}(t_{n,i}) - \tilde{Z}(t_{n,i})] = 0,$$

where $w_i \neq 0$. This result can be used in (2.5.1.8) to produce

$$\tilde{Y}(t_n^+) = \tilde{Y}_n + h_n w_0[-a(t_{n,0})\tilde{Y}(t_{n,0}) + \tilde{Z}(t_{n,0})] + h_n \sum_{j=1}^{m} w_0 L_j(c_0)[-a(t_{n,j})\tilde{Y}(t_{n,j}) + \tilde{Z}(t_{n,j})]. \tag{2.5.1.9}$$

The identity (2.5.1.7) allows us to write

$$\tilde{Y}(t_{n,i}) = \tilde{Y}(t_n^+) + h_n \sum_{j=1}^{m} \beta_j(c_i)[-a(t_{n,j})\tilde{Y}(t_{n,j}) + \tilde{Z}(t_{n,j})], \tag{2.5.1.10}$$

with

$$\beta_j(v) := \int_0^v L_j(s)ds \ (j = 1, \ldots, m),$$

and $\beta_j(c_i) =: a_{i,j}$. Hence, setting $\tilde{Y}_{n,i} := \tilde{Y}(t_{n,i})$ and recalling (2.5.1.9) we obtain

$$\begin{aligned} \tilde{Y}_{n,i} &= \tilde{Y}_n + h_n w_0[-a(t_{n,0})\tilde{Y}(t_{n,0}) + \tilde{Z}(t_{n,0})] \\ &+ h_n \sum_{j=1}^{m}[-a_{i,j} + w_0 L_j(c_0)][a(t_{n,j})\tilde{Y}(t_{n,j}) - \tilde{Z}(t_{n,j})] \end{aligned} \tag{2.5.1.11}$$

$(i = 1, \ldots, m)$. The equations (2.5.1.9) and (2.5.1.11) form a system of $m+1$ nonlinear algebraic equations for $\tilde{\mathbf{Y}}_n := (\ \tilde{Y}(t_n^+), \tilde{Y}_{n,1}, \ldots, \tilde{Y}_{n,m}\ )^T \in \mathbb{R}^{m+1}$: its form closely resembles the one corresponding to collocation at the points $\{t_{n,0}, t_{n,1}, \ldots, t_{n,m}\}$. We now show that these equations may indeed be interpreted as the stage equations of an implicit $(m+1)$-stage Volterra-Runge-Kutta (VRK) method. Let $b_j := \beta_j(1)$ $(j = 1, \ldots, m)$, and observe that

$$b_j = \int_0^1 L_j(s)ds = \sum_{k=0}^{m} w_k L_j(c_k) = w_0 L_j(c_0) + w_j,$$

because our interpolatory $(m+1)$-point quadrature formula is exact for polynomials of degree not exceeding m. This leads to the relationship

$$b_j - w_0 L_j(c_0) = w_j,$$

and hence by (2.5.1.7) to

$$\tilde{Y}_{n+1} := \tilde{Y}(t_{n+1}^-) = \tilde{Y}_n + h_n \sum_{j=0}^{m} w_j[-a(t_{n,j})\tilde{Y}(t_{n,j}) + \tilde{Z}(t_{n,j})]. \tag{2.5.1.12}$$

We conclude that (2.5.1.12) together with (2.5.1.9) and (2.5.1.10) represents a *collocation-based* $(m+1)$*-stage implicit VRK method* for (2.1.1.1). We summarize the above presentation as the following theorem.

Theorem 2.5.1.1. *The fully discretized DG scheme (2.5.1.4) may lead to the collocation-based* $(m+1)$*-stage implicit VRK method {(2.5.1.12), (2.5.1.9), (2.5.1.10)} for (2.1.1.1).*

Remark 2.5.1.1. *We see that the discussion is exactly the same in Section 1.3 for* $k \equiv 0$.

2.5.2 A posteriori error estimator

If the memory term $\mathcal{V}(y)(t)$ in (2.1.1.1) is computed approximately, then the resulting quadrature error also contributes to the total error. We consider the quadrature error as the perturbation of the DG(m) approximation to (2.1.1.1). The total error can be estimated by using the triangle inequality.

DG approximation to (2.1.1.1) is described by: Find $Y \in \mathcal{V}_N^{(m)}$ such that

$$B_{DG}(Y,X) = F_{DG}(X), \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.5.2.1}$$

where

$$\begin{aligned} B_{DG}(Y,X) &:= \sum_{n=1}^{M} \int_{I_n} \{Y'(t)X(t) + a(t)Y(t)X(t) - \mathcal{V}(Y)(t)X(t)\}dt \\ &+ \sum_{n=1}^{M-1} [Y]_n X_n^+ + Y_0^+ X_0^+, \\ F_{DG}(X) &:= y_0 X_0^+. \end{aligned}$$

Thus we define the discretized DG(m) with quadrature for the memory term as: Find $\tilde{Y} \in \mathcal{V}_N^{(m)}$,

$$\tilde{B}_{DG}(\tilde{Y}, X) = \tilde{F}_{DG}(X), \quad \forall X \in \mathcal{V}_N^{(m)}, \tag{2.5.2.2}$$

where

$$\begin{aligned}
\tilde{B}_{DG}(\tilde{Y}, X) &:= \sum_{n=1}^{M} \int_{I_n} \{\tilde{Y}'(t)X(t) + a(t)\tilde{Y}(t)X(t) - \tilde{V}(\tilde{Y})(t)X(t)\}dt \\
&+ \sum_{n=1}^{M-1} [\tilde{Y}]_n X_n^+ + \tilde{Y}_0^+ X_0^+, \\
\tilde{F}_{DG}(X) &:= y_0 X_0^+.
\end{aligned}$$

Thus the total error of the DG approximation to (2.1.1.1) with quadrature for the memory term can be written as

$$\tilde{e} := \tilde{Y} - y = (Y - y) + (\tilde{Y} - Y) := e + Q. \tag{2.5.2.3}$$

We analyze $e := Y - y$ as in the discussion in the above sections. The remaining work consists in estimating the term $Q := \tilde{Y} - Y$. By subtracting (2.5.2.1) from (2.5.2.2), we get

$$\tilde{B}_{DG}(\tilde{Y}, X) - B_{DG}(Y, X) = 0, \tag{2.5.2.4}$$

that is,

$$\begin{aligned}
&\sum_{n=1}^{M} \int_{I_n} \{Q'(t)X(t) + a(t)Q(t)X(t) - (\tilde{\mathcal{V}}(\tilde{Y})(t) - \mathcal{V}(Y(t)))X(t)\} \\
+ &\sum_{n=1}^{M-1} [Q]_n X_n^+ + Q_0^+ X_0^+ \\
= &\sum_{n=1}^{M} \int_{I_n} \{Q'(t)X(t) + a(t)Q(t)X(t) - \mathcal{V}(Q(t))X(t)\}dt \\
+ &\sum_{n=1}^{M-1} [Q]_n X_n^+ + Q_0^+ X_0^+ - \sum_{n=1}^{M} \int_{I_n} (\tilde{\mathcal{V}}(\tilde{Y})(t) - \mathcal{V}(\tilde{Y})(t))X(t)dt \\
=: \; & I + II = 0,
\end{aligned} \tag{2.5.2.5}$$

where

$$\begin{aligned} I &:= \sum_{n=1}^{M} \int_{I_n} \{Q'(t)X(t) + a(t)Q(t)X(t) - \mathcal{V}(Q(t))X(t)\}dt \\ &+ \sum_{n=1}^{M-1} [Q]_n X_n^+ + Q_0^+ X_0^+, \\ II &:= \sum_{n=1}^{M} \int_{I_n} (-\tilde{\mathcal{V}}(\tilde{Y})(t) + \mathcal{V}(\tilde{Y})(t))X(t)dt. \end{aligned}$$

These observations allow us to establish the following theorem. It focuses on VIDEs with *completely monotonic kernels*, due to their importance in many applications (see e.g., Gripenberg, Londen and Staffans [53, Ch. 5]).

Theorem 2.5.2.1. *Consider the discretized DG(m) with quadrature for the memory term ((2.5.2.2) with* $m = 0$ *and 1) for equation (2.1.1.1). We suppose that* $k \in C(\bar{I})$, $k \in C^d(R_+)$, *and k is completely monotonic:* $(-1)^j k^{(j)}(t) \geq 0$ $(t \geq 0,\ 0 \leq j \leq d)$, *and we take the quadrature form as*

$$\begin{aligned} \tilde{\mathcal{V}}(\tilde{Y})(t) &= \sum_{i=0}^{M-1} w_{Mi} k(t - t_i)\tilde{Y}(t_i^-) \\ &+ \sum_{i=1}^{M-1} \overline{w}_{Mi} k(t - t_i)\tilde{Y}(t_i^+) + w_{MM} k(0) t\tilde{Y}(t). \end{aligned} \tag{2.5.2.6}$$

Then

$$|\tilde{e}_M^-| := |\tilde{Y}_M^- - y(t_M)| \leq C||h_n^{m+1} R(\tilde{Y})||_{[0,t_M]}, \tag{2.5.2.7}$$

where $C := C(t_M, \mathcal{A}, \mathcal{B})$ *and* $\mathcal{A}$, $\mathcal{B}$ *are as in Theorem 2.1.2.1, and*

$$R(\tilde{Y}) := \frac{|[\tilde{Y}]_{n-1}|}{h_n} + |a(t)\tilde{Y} - \mathcal{V}(\tilde{Y})(t)|\ (t \in I_n).$$

Proof. First we prove (2.5.2.7) with $m = 0$. The well-known Peano theorem for quadrature [44] enables us to write the error as

$$
\begin{aligned}
\mathcal{E}(\tilde{Y})(t) &:= \mathcal{V}(\tilde{Y})(t) - \tilde{\mathcal{V}}(\tilde{Y})(t) \\
&= \sum_{i=0}^{M-1} \int_{I_i} K_q(s)\partial_s^{(q)}[k(t-s)\tilde{Y}(s)]ds \\
&+ \int_{t_{n-1}}^{t} K_q(s)\partial_s^{(q)}[k(t-s)\tilde{Y}(s)]ds, \qquad (2.5.2.8)
\end{aligned}
$$

where the Peano kernel is given by

$$
K_q(s) = \frac{1}{(q-1)!}\int_0^T (t-s)_+^{q-1}dt - \frac{1}{(q-1)!}\sum_{i=1}^{n} w_{ni}(t_i-s)_+^{q-1},
$$

with $q \geq 2$. From (2.5.2.5),

$$
\begin{aligned}
&\int_{I_n}\{Q'(t)X(t) + a(t)Q(t)X(t)\}dt + Q_{n-1}^{+}X_{n-1}^{+} \\
&= Q_{n-1}^{-}X_{n-1}^{+} - \int_{I_n}(\int_0^t k(t-s)Q(s)ds)X(t)dt \\
&- \int_{I_n}(\mathcal{V}(\tilde{Y})(t) - \tilde{\mathcal{V}}(\tilde{Y})(t))X(t)dt, \qquad (2.5.2.9)
\end{aligned}
$$

for all $X \in \mathcal{P}^{(m)}(I_n),\ n = 1, \cdots, M$. Selecting $X(t) = Q(t)$ in (2.5.2.9)

$$
\begin{aligned}
&\frac{1}{2}[Q_n^{-}]^2 + \frac{1}{2}[Q_{n-1}^{+}]^2 + \int_{I_n} aQ^2dt \\
&= Q_{n-1}^{-}Q_{n-1}^{+} - \int_{I_n}(\int_0^t k(t-s)Q(s)ds)Q(t)dt \\
&- \int_{I_n}\mathcal{E}(\tilde{Y})(t)X(t)dt. \qquad (2.5.2.10)
\end{aligned}
$$

Since $Q_{n-1}^{-}Q_{n-1}^{+} \leq \frac{1}{2}[Q_{n-1}^{-}]^2 + \frac{1}{2}[Q_{n-1}^{+}]^2$, we obtain

$$
\begin{aligned}
&\frac{1}{2}[Q_n^{-}]^2 + \int_{I_n} aQ^2dt \\
&\leq \frac{1}{2}[Q_{n-1}^{-}]^2 + \int_{I_n}(\int_0^t |k(t-s)||Q(s)|ds)|Q(t)|dt \\
&+ \int_{I_n}|\mathcal{E}(\tilde{Y})(t)||Q(t)|dt, \qquad (2.5.2.11)
\end{aligned}
$$

We notice that for $m = 0$,

$$\begin{aligned}
& \frac{1}{2}[Q_n^-]^2 + \int_{I_n} a(t)dt[Q_n^-]^2 \\
\leq \; & \frac{1}{2}[Q_{n-1}^-]^2 + \int_{I_n}\int_0^{t_{n-1}} |k(t-s)|dsdt \sum_{i=1}^{n-1} |Q_i^-||Q_n^-| \\
+ \; & \frac{h_n}{2}[Q_n^-]^2 + \frac{1}{2}\int_{I_n} [\mathcal{E}(\tilde{Y})(t)]^2 dt, \qquad (2.5.2.12)
\end{aligned}$$

we obtain, if h_n is sufficiently small, $0 < \beta_1 \leq \frac{1}{2}$,

$$\begin{aligned}
& [\frac{1}{2} - \frac{h_n}{2} - \frac{3}{2}\int_{I_n}\int_0^t |k(t-s)|dsdt][Q_n^-]^2 \\
\leq \; & \frac{1}{2}[Q_{n-1}^-]^2 + \frac{1}{2}\sum_{i=1}^{n-2}(\int_{I_n}\int_0^{t_{n-1}} |k(t-s)dsdt)|Q_i^-|^2 \\
+ \; & \frac{1}{2}\int_{I_n} [\mathcal{E}(\tilde{Y})(t)]^2 dt. \qquad (2.5.2.13)
\end{aligned}$$

We abbreviate (2.5.2.13) as

$$\beta_1[Q_n^-]^2 \leq \beta_2[Q_{n-1}^-]^2 + \frac{1}{2}\sum_{i=1}^{n-2}(\int_{I_n}\int_0^{t_{n-1}} |k(t-s)dsdt)|Q_i^-|^2 + \frac{1}{2}\int_{I_n}[\mathcal{E}(\tilde{Y})(t)]^2 dt,$$

where β_1, β_2 are obvious. Using discrete Gronwall lemma [26], we get

$$[Q_n^-]^2 \leq C\int_{I_n} (\mathcal{E}(\tilde{Y})(t))^2 dt. \qquad (2.5.2.14)$$

The estimate

$$|\tilde{e}_M^-| \leq |e_M^-| + |Q_M^-| \leq C(t_M, \mathcal{A}, \mathcal{B})||h_n R(Y)||_{[0,t_M]} + |Q_M^-| \qquad (2.5.2.15)$$

now follows from (2.5.2.3) and Theorem 2.1.2.1. Because $Y = \tilde{Y} - Q$,

$$\begin{aligned}
||h_n R(Y)||_{[0,t_M]} &= \max_{1\leq n\leq M} ||h_n R(Y)||_{\bar{I}_n} \\
&= \max_{1\leq n\leq M} \{h_n \frac{|[\tilde{Y}]_{n-1} - [Q]_{n-1}|}{h_n} \\
&+ h_n ||a(t)\tilde{Y} - a(t)Q_n^- - \int_0^t k(t-s)[\tilde{Y} - Q_n^-]ds||_{\bar{I}_n}\} \\
&\leq \max_{1\leq n\leq M} \{h_n \frac{|[\tilde{Y}]_{n-1}| + |[Q]_{n-1}|}{h_n} \\
&+ h_n ||a(t)\tilde{Y} - \int_0^t k(t-s)\tilde{Y}(s)ds||_{\bar{I}_n} \\
&+ h_n ||a(t)Q_n^-||_{\bar{I}_n} + h_n || \int_0^t k(t-s)Q_n^- ds||_{\bar{I}_n}\} \\
&\leq |h_n R(\tilde{Y})| + \max_{1\leq n\leq M} \{|[Q]_{n-1}| \\
&+ ||a(t)Q_n^-||_{\bar{I}_n} + || \int_0^t k(t-s)Q_n^- ds||_{\bar{I}_n}\}, \qquad (2.5.2.16)
\end{aligned}$$

Combining (2.5.2.14), (2.5.2.15), (2.5.2.16), and noting that

$$\max_{1\leq n\leq M} \{|[Q]_{n-1}| + ||a(t)Q_n^-||_{\bar{I}_n} + || \int_0^t k(t-s)Q_n^- ds||_{\bar{I}_n}\} = \mathcal{O}(h^q),$$

with $q \geq 2$, we obtain (2.5.2.7) with $m = 0$.

We shall now prove (2.5.2.7) with $\underline{m = 1}$. Recall (2.5.2.10):

$$\begin{aligned}
&\frac{1}{2}[Q_n^-]^2 + \frac{1}{2}[Q_{n-1}^+]^2 + \int_{I_n} aQ^2 dt \\
&= Q_{n-1}^- Q_{n-1}^+ - \int_{I_n} (\int_0^t k(t-s)Q(s)ds)Q(t)dt \\
&- \int_{I_n} \mathcal{E}(\tilde{Y})(t)Q(t)dt. \qquad (2.5.2.17)
\end{aligned}$$

Since $Q_{n-1}^- Q_{n-1}^+ \leq \epsilon^2 [Q_{n-1}^-]^2 + \frac{1}{\epsilon^2}[Q_{n-1}^+]^2$ and for $m = 1$, $||Q||_{\bar{I}_n}^2 \leq |Q_{n-1}^+|^2 + |Q_n^-|^2$, we obtain

$$\begin{aligned} [Q_n^-]^2 + [Q_{n-1}^+]^2 &\leq C \int_{I_n} \left(\mathcal{E}(\tilde{Y})\right)^2 dt + C[Q_{n-1}^-]^2 \\ &+ C \sum_{i=1}^{n-2} h_i \left([Q_i^-]^2 + [Q_i^+]^2\right). \end{aligned} \tag{2.5.2.18}$$

Using again discrete Gronwall lemma [26], we reach

$$[Q_n^-]^2 + [Q_{n-1}^+]^2 \leq C \int_{I_n} \left(\mathcal{E}(\tilde{Y})\right)^2 dt. \tag{2.5.2.19}$$

Because $Y = \tilde{Y} - Q$,

$$\begin{aligned} ||h_n^2 R(Y)||_{[0,t_M]} &= \max_{1\leq n\leq M} ||h_n^2 R(Y)||_{\bar{I}_n} \\ &= \max_{1\leq n\leq M} \{h_n^2 \frac{|[\tilde{Y}]_{n-1} - [Q]_{n-1}|}{h_n} \\ &+ h_n^2 ||a(t)\tilde{Y} - a(t)Q - \int_0^t k(t-s)[\tilde{Y} - Q]ds||_{\bar{I}_n}\} \\ &\leq \max_{1\leq n\leq M} \{h_n^2 \frac{|[\tilde{Y}]_{n-1}| + |[Q]_{n-1}|}{h_n} \\ &+ h_n^2 ||a(t)\tilde{Y} - \int_0^t k(t-s)\tilde{Y}(s)ds||_{\bar{I}_n} \\ &+ h_n^2 ||a(t)Q||_{\bar{I}_n} + h_n^2 || \int_0^t k(t-s)Qds||_{\bar{I}_n}\} \\ &\leq |h_n^2 R(\tilde{Y})| + \max_{1\leq n\leq M} \{h_n |[Q]_{n-1}| \\ &+ h_n^2 ||a(t)Q||_{\bar{I}_n} + h_n^2 || \int_0^t k(t-s)Qds||_{\bar{I}_n}\}. \end{aligned} \tag{2.5.2.20}$$

Combining (2.5.2.19), (2.5.2.15), (2.5.2.20), an observing that

$$\max_{1\leq n\leq M} \{|[Q]_{n-1}| + ||a(t)Q||_{\bar{I}_n} + || \int_0^t k(t-s)Qds||_{\bar{I}_n}\} = \mathcal{O}(h^q),$$

with $q \geq 2$, we obtain (2.5.2.7) with $m = 1$.

2.6 Superconvergence of mesh-dependent Galerkin methods for VIDEs

In this section we extend the mesh-dependent Galerkin methods (including the discontinuous Galerkin method) of Section 1.2 to Volterra integro-differential equations, including the semilinear case. Furthermore, theorems on superconvergence are proven.

2.6.1 Superconvergence for semilinear VIDEs

Consider again the semilinear Volterra integro-differential equation

$$\begin{cases} y'(t) + a(t)y(t) = \mathcal{V}_G(y)(t), & t \in I = [0,T], \\ y(0) = y_0, \end{cases} \tag{2.6.1.1}$$

where $\mathcal{V}_G(y)(t) := \displaystyle\int_0^t k(t-s)G(y)(s)ds$. Assume that $a,\ k \in C(I)$, and G is Lipschitz continuous, i.e.,

$$|G(y_1) - G(y_2)| \leq L|y_1 - y_2|, \tag{2.6.1.2}$$

for all $y_1,\ y_2 \in \Omega \subset \mathbb{R}$.

As in Section 1.2.1, we introduce the mesh-dependent weak form of (2.6.1.1): Find

$$\tilde{u} := (U_0, \dots, U_N, u_1, \dots, u_N) \in \mathcal{U} := \Omega^{N+1} \times \prod_{n=1}^{N} L^2(I_n; \Omega),$$

such that

$$\begin{aligned} & U_0[V_0 - v_1(t_0)] + \sum_{n=1}^{N-1} U_n[v_n(t_n) - v_{n+1}(t_n)] + U_N v_N(t_N) \\ - & \sum_{n=1}^{N} \int_{I_n} u_n v_n' dt = y_0 V_0 + \sum_{n=1}^{N} \int_{I_n} [-a(t)u_n \\ + & \int_0^t k(t-s)G(u(s))ds] v_n dt, \end{aligned} \tag{2.6.1.3}$$

for all

$$\tilde{v} := (V_0, v_1, \ldots, v_N) \in \mathcal{V} := \Omega \times \Pi_{n=1}^{N} H^1(I_n; \Omega),$$

where $u = \sum_{i=1}^{n} u_i \chi_{I_i}$, with χ_{I_i} denoting the characteristic function of I_i.

The local meaning of (2.6.1.3) is to find u_n in $L^2(I_n; \Omega)$ and U_n in Ω such that

$$\begin{cases} U_n v_n(t_n) = U_{n-1} v_n(t_{n-1}) + \int_{I_n} [u_n v_n' \\ +(-a(t)u_n(t) + \int_0^t k(t-s)G(u(s))ds)v_n]dt, \\ U_0 = y_0, \end{cases} \tag{2.6.1.4}$$

for all $v_n \in H^1(I_n; \Omega)$ and $n = 1, \ldots, N$.

Thus the mesh-dependent Galerkin scheme for (2.6.1.4) is to find $\tilde{u}_h \in \mathcal{U}_h$ such that $U_0 = y_0$ and

$$\begin{cases} U_n^h v_n^h(t_n) - \int_{I_n} u_n^h (v_n^h)' dt = U_{n-1}^h v_n^h(t_{n-1}) + \int_{I_n} [-a(t)u_n^h \\ + \int_0^t k(t-s)G(u^h(s))ds] v_n^h dt, \\ J \text{ additional conditions on } u_n^h, \end{cases} \tag{2.6.1.5}$$

for all v_n^h in $\mathcal{P}^{(m+1-J)}(I_n; \Omega)$ and $n = 1, \ldots, N$. Here $u^h = \sum_{i=1}^{n-1} u_i^h \chi_i$ and

$$\mathcal{U}_h = \left\{ \tilde{u}_h \;\middle|\; \begin{array}{l} \tilde{u}_h = (U_0^h, \ldots, U_N^h, u_1^h, \ldots, u_N^h) \in \mathcal{U}_h \text{ such that } u_n^h \in \mathcal{P}^{(m)}(I_n; \Omega) \\ \text{subject to } J \ (\geq 0) \text{ additional conditions for } n = 1, \ldots, N. \end{array} \right\},$$

We estimate the L^2- and nodal error of the mesh-dependent Galerkin scheme (2.6.1.5) for VIDE (2.6.1.1) in the following theorem.

Theorem 2.6.1.1. *Assume that the solution y of (2.6.1.1) belongs to $H^{m+1}([0,T]; \Omega)$. For $M > 1$, assume that on the first $M-1$ intervals the solution of (2.6.1.5) is such that*

$$\max\{|U_n^h - y(t_n)| : \ n = 0, \ldots, M-1\} \leq Ch^{m+1} ||y^{(m+1)}||_0, \tag{2.6.1.6}$$

and for $j = 0, \ldots, m+1$,

$$\left\{ \sum_{n=1}^{M-1} ||u_n^h - y||_{j,n}^2 \right\}^{1/2} \leq Ch^{m+1-j}||y^{(m+1)}||_0. \qquad (2.6.1.7)$$

Subsequently, we have that for sufficiently small h,

$$\max\{|U_n^h - y(t_n)| : \ n = 0, \ldots, N\} \leq Ch^{m+1}||y^{(m+1)}||_0, \qquad (2.6.1.8)$$

and for $j = 0, \ldots, m+1$,

$$||u^h - y||_j \leq Ch^{m+1-j}||y^{(m+1)}||_0, \qquad (2.6.1.9)$$

where $u_h = \sum_{n=1}^{N} u_n^h \chi_{I_n}$ *and* $|| \cdot ||_j := \left\{ \sum_{n=1}^{N} || \cdot ||_{j,n}^2 \right\}^{1/2}$.

Proof. Since y solves (2.6.1.5), we have

$$\begin{aligned} (U_n^h - y(t_n))v_n^h(t_n) &= [U_{n-1}^h - y(t_{n-1})]v_n^h(t_{n-1}) + \int_{I_n} (u_n^h - y)(v_n^h)' dt \\ &- \int_{I_n} a(t)(u_n^h - y)v_n^h dt + \int_{I_n} \int_0^t k(t-s)[G(u^h(s)) \\ &- G(y(s))]dsv_n^h dt. \end{aligned} \qquad (2.6.1.10)$$

Let v_n^h be the solution of

$$(v_n^h)'(t) = -\wp_J(u_n^h - \bar{u}_n^h)(t), \quad t \in I_n, \quad v_n^h(t_n) = 0,$$

where $\bar{u}_n^h$ is the Lagrange interpolating polynomial of degree m, such that

$$\bar{u}_n^h(t_{n_l}) = y(t_{n_l}), \quad l = 1, \ldots, J.$$

$\wp_J$ denotes the L^2-projector of $L^2(I_n;\Omega)$ onto $\mathcal{P}^{(m-J)}(I_n;\Omega)$. We substitute v_n^h into

(2.6.1.10) and obtain

$$
\begin{aligned}
&\int_{I_n} (u_n^h - \bar{u}_n^h)[\wp_J(u_n^h - \bar{u}_n^h)(t)]dt \\
= \;& [U_{n-1}^h - y(t_{n-1})] \int_{I_n} \wp_J(u_n^h - \bar{u}_n^h)(t)dt + \int_{I_n} (\bar{u}_n^h - y)[-\wp_J(u_n^h - \bar{u}_n^h)(t)]dt \\
- \;& \int_{I_n} a(t)(u_n^h - y) \int_t^{t_n} \wp_J(u_n^h - \bar{u}_n^h)(\nu)d\nu dt + \int_{I_n} \{ \int_0^t k(t-s)[G(u^h(s)) \\
- \;& G(y(s))]ds \int_t^{t_n} \wp_J(u_n^h - \bar{u}_n^h)(\nu)d\nu \}dt.
\end{aligned}
$$

Hence we arrive at

$$
\begin{aligned}
\|\wp_J(u_n^h - \bar{u}_n^h)(t)\|_{0,n} \;\le\;& h_n^{1/2}|U_{n-1} - y(t_{n-1})| + h_n^{1/2}\|a(t)\|_{0,n}\|u_n^h - \bar{u}_n^h\|_{0,n} \\
+\;& [1 + h_n^{1/2}\|a(t)\|_{0,n}]\|\bar{u}_n^h - y\|_{0,n} \\
+\;& Lh_n^{1/2} \int_{I_n} \int_0^t |k(t-s)||u^h(s) - y(s)|dsdt. \qquad (2.6.1.11)
\end{aligned}
$$

We need Lemma 2.6.1.1 whose proof can be found in Delfour and Dubeau [40].

Lemma 2.6.1.1. *The map J_n defined by*

$$
u \to J_n u = (\wp_J u, u(t_{n_1}), \ldots, u(t_{n_J})) : \; \mathcal{P}^{(m)}(I_n; \Omega) \to \mathcal{P}^{(m-J)}(I_n; \Omega) \times \Omega^J
$$

is an isomorphism, and there exist two constants β_1 and β_2 (independent of h and the points $\{t_{n_l}\}_{l=1}^J$) such that

$$
\beta_1 \|u\|_{0,n} \le \|J_n u\| \le \beta_2 \|u\|_{0,n}.
$$

It follows from Lemma 2.6.1.1 and (2.6.1.11) that

$$
\begin{aligned}
& [\beta_1 - h_n^{1/2}\|a(t)\|_{0,n} - Lh_n^{3/2}\|k(t)\|_{0,n}]\|u_n^h - \bar{u}_n^h\|_{0,n} \\
\le\;& h_n^{1/2}|U_{n-1}^h - y(t_{n-1})| + [1 + h_n^{1/2}\|a(t)\|_{0,n}]\|\bar{u}_n^h - y\|_{0,n} \\
+\;& Lh_n^{3/2}\|k(t)\|_{0,n}\|\bar{u}_n^h(t) - y(t)\|_{0,n} + h_n^{1/2} \sum_{l=1}^{J} |u_n^h(t_{n_l}) - \bar{u}_n^h(t_{n_l})| \\
+\;& Lh_n^{1/2} \int_{I_n} \sum_{j=1}^{n-1} \int_{I_j} |k(t-s)||u_j^h(s) - y(s)|dsdt
\end{aligned}
$$

Hence,

$$
\begin{aligned}
||u_n^h - y||_{0,n} &\leq ||u_n^h - \bar{u}_n^h||_{0,n} + ||\bar{u}_n^h - y||_{0,n} \\
&\leq Ch_n^{1/2}\{|U_{n-1}^h - y(t_{n-1})| + \sum_{l=1}^{J} |u_n^h(t_{n_l}) - \bar{u}_n^h(t_{n_l})|\} \\
&+ C||\bar{u}_n^h(t) - y(t)||_{0,n} \\
&+ Ch_n^{1/2} \int_{I_n} \sum_{j=1}^{n-1} \int_{I_j} |k(t-s)||u_j^h(s) - y(s)|dsdt. \qquad (2.6.1.12)
\end{aligned}
$$

Substitute $v_n^h = U_n^h - y(t_n)$ into (2.6.1.10):

$$
\begin{aligned}
|U_n^h - y(t_n)| &\leq |U_{n-1}^h - y(t_{n-1})| + ||a(t)||_{0,n}||u_n^h - y||_{0,n} \qquad (2.6.1.13) \\
&+ Lh_n^{3/2}||k(t)||_{0,n}||u_n^h(t) - y(t)||_{0,n} \\
&+ Lh_n^{1/2} \int_{I_n} \left(\sum_{j=1}^{n-1} \int_{I_j} |k(t-s)||u_j^h(s) - y(s)|ds \right) dt.
\end{aligned}
$$

We therefore obtain

$$
\begin{aligned}
|U_n^h - y(t_n)| &\leq \sum_{i=1}^{n} ||a(t)||_{0,i}||u_i^h - y||_{0,i} + \sum_{i=1}^{n} Lh_i^{3/2}||k(t)||_{0,i}||u_i^h - y||_{0,i} \\
&+ \sum_{i=1}^{n} Lh_i^{1/2} \int_{I_i} \sum_{j=1}^{n-1} \int_{I_j} |k(t-s)||u_j^h(s) - y(s)|dsdt. \qquad (2.6.1.14)
\end{aligned}
$$

We note that for $J \leq M$ and $u_n^h(t_{n_l}) = U_{n_l}^h$, inequality (2.6.1.12) can be rearranged to read

$$
\begin{aligned}
||u_n^h - y||_{0,n} &\leq Ch_n^{1/2} \sum_{i=1}^{M} |U_{n-i}^h - y(t_{n-i})| + Ch^{m+1}||y^{(m+1)}||_{0,\bar{n}} \\
&+ Ch_n^{1/2} \int_{I_n} \sum_{j=1}^{n-1} \int_{I_i} |k(t-s)||u_j^h(s) - y(s)|dsdt, \qquad (2.6.1.15)
\end{aligned}
$$

where $||\cdot||_{0,\bar{n}}$ is the L^2-norm over $[t_{n-M}, t_n]$. Set $\alpha_n := |U_n^h - y(t_n)|, \quad n = 1, \ldots, N,$ and

$$
\begin{aligned}
\beta_j &:= ||a(t)||_{0,j}||u_j^h - y||_{0,j} + Lh_j||k(t)||_{0,j}||u_j^h - y||_{0,j} \\
&+ Lh_j^{1/2} \int_{I_j} \sum_{m=1}^{n-1} \int_{I_m} |k(t-s)||u_m^h(s) - y(s)|dsdt,
\end{aligned}
$$

for $j = 1, \ldots, N$. If we abbreviate (2.6.1.14) by setting $\alpha_n \leq \sum_{j=1}^n \beta_j$, then

$$\sum_{i=1}^{M} |U_{n-i}^h - y(t_{n-i})| = \sum_{i=1}^{M} \alpha_{n-i} \leq \sum_{i=1}^{M} \sum_{j=1}^{n-i} \beta_j \leq M \sum_{j=1}^{n-1} \beta_j.$$

Since

$$\begin{aligned} \beta_j &\leq h_j^{1/2} \mathcal{A} ||u_j^h - y||_{0,j} + L h_j^{3/2} \mathcal{B} ||u_j^h - y||_{0,j} \\ &+ L\mathcal{B} h_j^{3/2} \sum_{m=1}^{n-1} h_m^{1/2} ||u_m^h(s) - y(s)||_{0,m}, \end{aligned}$$

we have

$$\sum_{i=1}^{M} ||U_{n-i}^h - y(t_{n-i})| \leq L\mathcal{B}TM \sum_{m=1}^{n-1} (h_m^{1/2} + \frac{\mathcal{A} h_m^{1/2}}{L\mathcal{B}T} + \frac{h_m^{3/2}}{T}) ||u_m^h - y||_{0,m}.$$

Recalling (2.6.1.15), we then derive the bound

$$\begin{aligned} ||u_n^h - y||_{0,n} &\leq C h_n^{1/2} L\mathcal{B}TM \sum_{m=1}^{n-1} (h_m^{1/2} + \frac{\mathcal{A} h_m^{1/2}}{L\mathcal{B}T} + \frac{h_m^{3/2}}{T}) ||u_m^h - y||_{0,m} \\ &+ C\mathcal{B} h_n \sum_{m=1}^{n-1} h_m^{1/2} ||u_m^h - y||_{0,m} + C h^{m+1} ||y^{(m+1)}||_{0,\bar{n}} \\ &\leq C h_n^{1/2} \sum_{m=1}^{n-1} h_m^{1/2} ||u_m^h - y||_{0,m} + C h^{m+1} ||y^{(m+1)}||_{0,\bar{n}}, \end{aligned}$$

and Gronwall's lemma leads to

$$\begin{aligned} ||u_n^h - y||_{0,n} &\leq C h^{m+1} \exp(C h_n^{1/2} \sum_{m=1}^{n-1} h_m^{1/2}) ||y^{(m+1)}||_{0,\bar{n}} \\ &\leq C h^{m+1} ||y^{(m+1)}||_{0,\bar{n}}. \end{aligned}$$

Thus (2.6.1.7) and (2.6.1.9) hold true when $j = 0$. From (2.6.1.13) and Gronwall's lemma we also obtain (2.6.1.6) and (2.6.1.8). The inequalities (2.6.1.9) ($1 \leq j \leq m+1$) are obtained by using the estimates

$$||u_n^h - y||_{j,n} \leq ||u_n^h - \bar{u}_n^h||_{j,n} + ||\bar{u}_n^h - y||_{j,n}$$

and

$$||u_n^h - \bar{u}_n^h||_{j,n} \leq Ch^{-j}||u_n^h - \bar{u}_n^h||_{0,n}.$$

Now we describe the superconvergence property of the mesh-dependent Galerkin method (2.6.1.5) for the VIDE (2.6.1.1).

Theorem 2.6.1.2. *Assume that the assumptions of Theorem 2.6.1.1 hold. Then for all sufficently small $h > 0$ we have*

$$\max\{|U_n^h - y(t_n)| : \ n = 0, \dots, N\} \leq Ch^{2m+2-J}. \tag{2.6.1.16}$$

Proof. Choose $\tilde{v}^h = (V_0, v_1^h, \dots, v_N^h)$ such that $v_1^h(t_0) = V_0$ and $v_n^h(t_n) = v_{n+1}^h(t_n)$, $n = 1, \dots, N-1$. Substitute that $\tilde{v}^h$ in (2.6.1.10) and sum over $j = 1, \dots, n$. This yields

$$\begin{aligned}(U_n^h - y(t_n))v_n^h(t_n) &= \int_0^{t_n} (u^h - y)(v^h)' dt - \int_0^{t_n} a(t)(u^h - y)v^h dt \\ &+ \int_0^{t_n}\int_0^t k(t-s)[G(u^h(s)) \\ &- G(y(s))]ds\, v^h dt. \end{aligned} \tag{2.6.1.17}$$

Let w in $H^{m+2-J}(0, t; \Omega)$ be the solution of

$$\begin{cases} w' - a(t)w + \int_t^{t_n} k(s-t) \int_0^1 G_1(ru^h + (1-r)y)dr \cdot w(s)ds = 0, \\ w(t_n) = U_n^h - y(t_n), \end{cases}$$

for $t \in [0, t_n]$. Let w^h be a continuous piecewise interpolating polynomial of degree $m+1-J$ of w such that $w^h(t_n) = U_n^h - y(t_n)$. It follows from Sobolev interpolation theory [22] that

$$||w - w^h||_1 \leq Ch^{m+1-J}||w^{(m+2-J)}||_0.$$

Now we shall use the following Lemma 2.6.1.2 to express the norm of $w^{(m+2-J)}$ in terms of $U_n^h - y(t_n)$.

Lemma 2.6.1.2. *Suppose that g is such that $g|_{I_n} \in H^{m+1-J}(I_n;\Omega)$ $(n = 1, \ldots, N)$, and that w in $H^1([0,T];\Omega)$ satisfies the equation*

$$w'(t) - a(t)w(t) + \int_t^{t_n} k(s-t)\tilde{A}(s)w(s)ds = g(t), \tag{2.6.1.18}$$

for $t \in [0, t_n]$. Then there exists a constant C, independent of g and t_n, such that

$$||w||_{m+2-J} \leq C\{|w(t_n)| + ||g||_{m+1-J}\}.$$

Proof. We differentiate (2.6.1.18) $m+1-J$ times to express $w^{(m+2-J)}$ in terms of $\{w,\ \mathcal{V}_q^* w(t)\ (q = 0, \ldots, m+1-J),\ g,\ g^{(1)}, \ldots, g^{(m+1-J)}\}$, where

$$\mathcal{V}_q^* w(t) := \int_t^{t_n} k_t^{(q)}(s-t)\tilde{A}(s)w(s)ds.$$

Then we replace w by the identity

$$w(t) = R(t, t_n)w(t_n) + \int_t^{t_n} R(t,s)g(s)ds, \tag{2.6.1.19}$$

where the resolvent kernel R has the form

$$R(t,s) = 1 + \int_t^s r(t,u)du, \quad (t,s) \in S := \{(t,s) :\ 0 \leq t \leq s \leq t_n \leq T\},$$

with r satisfying

$$r(t,s) = Q(t,s) + \int_t^s Q(t,\tau)r(\tau,s)d\tau,$$

and with

$$Q(t,s) := a(s) + \int_t^s k_t^{(q)}(u-s)du, \quad (t,s) \in S.$$

We refer to ([26] or [23]) for the proof of (2.6.1.19). The proof of Lemma 2.6.1.2 is now complete.

We then have

$$||w - w^h||_1 \leq Ch^{m+1-J}|U_n^h - y(t_n)|.$$

Set $v^h = w^h$ in (2.6.1.17):

$$\begin{aligned}
|U_n^h - y(t_n)|^2 &\leq ||u^h - y||_0 ||(w^h)' - a(t)w^h + \int_t^{t_n} k(s-t)\tilde{A}(s)w^h(s)ds||_0 \\
&\leq ||u^h - y||_0 ||(w^h)' - a(t)w^h + \int_t^{t_n} k(s-t)\tilde{A}(s)w^h(s)ds \\
&- [w' - a(t)w + \int_t^{t_n} k(s-t)\tilde{A}(s)w(s)ds]||_0 \\
&\leq ||u^h - y||_0 \{||(w^h)' - w'||_0 + \mathcal{A}||w^h - w||_0 + TL\mathcal{B}||w^h - w||_0\} \\
&\leq ||u^h - y||_0 \{Ch^{m+1-J}|U_n^h - y(t_n)| \\
&+ C(\mathcal{A} + TL\mathcal{B})h^{m+2-J}|U_n^h - y(t_n)|\}.
\end{aligned}$$

Hence,

$$|U_n^h - y(t_n)| \leq Ch^{m+1-J}||u^h - y||_0.$$

Combining this equation with (2.6.1.9) we arrive at the desired estimate (2.6.1.16).

2.6.2 Superconvergence of the discretized mesh-dependent Galerkin methods for VIDEs

Consider, for ease of exposition, the linear VIDE

$$\begin{cases} y'(t) + a(t)y(t) = \mathcal{V}(y)(t), & t \in I = [0,T], \\ y(0) = y_0, \end{cases} \tag{2.6.2.1}$$

where $\mathcal{V}(y)(t) := \int_0^t k(t-s)y(s)ds$ and $a,\ k \in C(I)$.

As in Section 1.2.1, we introduce the mesh-dependent weak form of (2.6.1.1): To find

$$\tilde{u} := (U_0, \ldots, U_N, u_1, \ldots, u_N) \in \mathcal{U} := \Omega^{N+1} \times \prod_{n=1}^{N} L^2(I_n; \Omega),$$

such that

$$\begin{aligned}
& U_0[V_0 - v_1(t_0)] + \sum_{n=1}^{N-1} U_n[v_n(t_n) - v_{n+1}(t_n)] + U_N v_N(t_N) \\
- \quad & \sum_{n=1}^{N} \int_{I_n} u_n v_n' dt = y_0 V_0 + \sum_{n=1}^{N} \int_{I_n} [-a(t) u_n \\
+ \quad & \mathcal{V}(u)(t)] v_n dt, \quad \forall \tilde{v} \in \mathcal{V},
\end{aligned} \tag{2.6.2.2}$$

where $u = \sum_{i=1}^{n} u_i \chi_{I_i}$.

The local form of (2.6.2.2) is: Find u_n in $L^2(I_n; \Omega)$ and U_n in Ω such that

$$\begin{cases} U_n v_n(t_n) = U_{n-1} v_n(t_{n-1}) + \displaystyle\int_{I_n} [u_n v_n' \\ +(-a(t) u_n(t) + \mathcal{V}(u)(t)) v_n] dt, \\ U_0 = y_0, \end{cases} \tag{2.6.2.3}$$

for all $v_n \in H^1(I_n; \Omega)$ and $n = 1, \ldots, N$.

Thus the approximation scheme for (2.6.2.3) consists in finding $\tilde{u}_h$ in $\mathcal{U}_h$ such that $U_0 = y_0$ and

$$\begin{cases} U_n^h v_n^h(t_n) - \displaystyle\int_{I_n} u_n^h (v_n^h)' dt = U_{n-1}^h v_n^h(t_{n-1}) + \int_{I_n} [-a(t) u_n^h \\ +\mathcal{V}(u^h(s))(t)] v_n^h dt, \\ J \text{ additional conditions on } u_n^h, \end{cases} \tag{2.6.2.4}$$

for all v_n^h in $\mathcal{P}^{(m+1-J)}(I_n; \Omega)$ and $n = 1, \ldots, N$. Here $u^h = \sum_{i=1}^{n-1} u_i^h \chi_i$ and

$$\mathcal{U}_h = \left\{ \tilde{u}_h \,\middle|\, \begin{array}{l} \tilde{u}_h = (U_0^h, \ldots, U_N^h, u_1^h, \ldots, u_N^h) \in \mathcal{U} \ \text{ such that } u_n^h \in \mathcal{P}^{(m)}(I_n; \Omega) \\ \text{subject to } J\ (\geq 0) \text{ additional conditions for } n = 1, \ldots, N. \end{array} \right\},$$

When we apply the numerical quadrature to the memory term, we obtain the semi-discretized DG scheme for (2.6.2.3): Find $\tilde{\tilde{u}}_h$ in $\tilde{\mathcal{U}}_h$ such that $\tilde{U}_0 = y_0$ and

$$\begin{cases} \tilde{U}_n^h \tilde{v}_n^h(t_n) - \displaystyle\int_{I_n} \tilde{u}_n^h (\tilde{v}_n^h)' dt = \tilde{U}_{n-1}^h \tilde{v}_n^h(t_{n-1}) + \int_{I_n} [-a(t) \tilde{u}_n^h \\ +\tilde{\mathcal{V}}(\tilde{u}^h)(t)] \tilde{v}_n^h dt, \\ J \text{ additional conditions on } \tilde{u}_n^h, \end{cases} \tag{2.6.2.5}$$

for all $\tilde{v}_n^h$ in $\mathcal{P}^{(m+1-J)}(I_n;\Omega)$ and $n = 1, \ldots, N$. Here $\tilde{u}^h = \sum_{i=1}^{n-1} \tilde{u}_i^h \chi_i$ and

$$\tilde{\mathcal{U}}_h = \left\{ \tilde{\tilde{u}}_h \;\middle|\; \begin{array}{l} \tilde{\tilde{u}}_h = (\tilde{U}_0^h, \ldots, \tilde{U}_N^h, \tilde{u}_1^h, \ldots, \tilde{u}_N^h) \in \tilde{\mathcal{U}} \text{ such that } \tilde{u}_n^h \in \mathcal{P}^{(m)}(I_n;\Omega) \\ \text{subject to } J\ (\geq 0) \text{ additional conditions for } n = 1, \ldots, N, \end{array} \right\},$$

where we will use the interpolatory quadrature approximation (for example, Newton-Cotes formulas [44]) for the memory term:

$$\tilde{\mathcal{V}}(\tilde{u}^h)(t) := \sum_{i=0}^{n-1} \omega_{ni} k(t_{n-1} - t_i) \tilde{U}_i^h + \omega_{nn} k(0) t \tilde{u}_n^h(t). \tag{2.6.2.6}$$

To make the error of the quadrature formula (2.6.2.6) be $\mathcal{O}(h^m)$ for all $t \in (0, T]$, we adapt the old mesh (with meshsize $h := \max\limits_{(n)}\{h_n,\ n = 1, \ldots, N\}$) by choosing

$$\max\{h_i :\ 1 \leq i \leq m\} := h^m.$$

We describe the L^2 and nodal error of the semi-discretized DG scheme (2.6.2.5) for (2.6.2.1) in the following theorem.

Theorem 2.6.2.1. *Assume that the solution y of (2.6.2.1) belongs to $H^{m+1}([0,T];\Omega)$. For $M > 1$, assume that on the first $M-1$ intervals the solution of (2.6.2.5) is such that*

$$\max\{|\tilde{U}_n^h - y(t_n)| :\ n = 0, \ldots, M-1\} \leq C h^{m+1} ||y^{(m+1)}||_0, \tag{2.6.2.7}$$

and for $j = 0, \ldots, m+1$,

$$\left\{ \sum_{n=1}^{M-1} ||\tilde{u}_n^h - y||_{j,n}^2 \right\}^{1/2} \leq C h^{m+1-j} ||y^{(m+1)}||_0. \tag{2.6.2.8}$$

Hence, we have that for sufficiently small $h > 0$,

$$\max\{|\tilde{U}_n^h - y(t_n)| :\ n = 0, \ldots, N\} \leq c h^{m+1} ||y^{(m+1)}||_0, \tag{2.6.2.9}$$

and for $j = 0, \ldots, m+1$,

$$||\tilde{u}^h - y||_j \leq C h^{m+1-j} ||y^{(m+1)}||_0, \tag{2.6.2.10}$$

where $\tilde{u}_h = \sum_{n=1}^{N} \tilde{u}_n^h \chi_{I_n}$ and $||\cdot||_j := \left\{ \sum_{n=1}^{N} ||\cdot||_{j,n}^2 \right\}^{1/2}$.

Proof. Since y satisfies (2.6.2.5), we have

$$\begin{aligned}(\tilde{U}_n^h - y(t_n))\tilde{v}_n^h(t_n) &= [\tilde{U}_{n-1}^h - y(t_{n-1})]\tilde{v}_n^h(t_{n-1}) + \int_{I_n}(\tilde{u}_n^h - y)(\tilde{v}_n^h)'dt \\ &- \int_{I_n} a(t)(\tilde{u}_n^h - y)\tilde{v}_n^h dt + \int_{I_n}[\tilde{\mathcal{V}}(\tilde{u}^h)(t) \\ &- \mathcal{V}(y)(t)]\tilde{v}_n^h dt. \qquad (2.6.2.11)\end{aligned}$$

Let $\tilde{v}_n^h$ be the solution of

$$(\tilde{v}_n^h)'(t) = -\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(t), \quad t \in I_n, \quad \tilde{v}_n^h(t_n) = 0,$$

where $\bar{u}_n^h$ is the Lagrange interpolating polynomial of degree m,

$$\overline{u}_n^h(t_{n_l}) = y(t_{n_l}), \quad l = 1, \ldots, J.$$

$\wp_J$ is the L^2-projector of $L^2(I_n;\Omega)$ onto $\mathcal{P}^{(m-J)}(I_n;\Omega)$. Then we substitute $\tilde{v}_n^h$ into (2.6.2.11) and obtain

$$\begin{aligned}&\int_{I_n}(\tilde{u}_n^h - \bar{u}_n^h)[\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(t)]dt \\ = \ & [\tilde{U}_{n-1}^h - y(t_{n-1})]\int_{I_n}\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(t)dt + \int_{I_n}(\bar{u}_n^h - y)[-\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(t)]dt \\ - \ & \int_{I_n} a(t)(\tilde{u}_n^h - y)\int_t^{t_n}\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(\nu)d\nu dt + \int_{I_n}[\tilde{\mathcal{V}}(\tilde{u}^h)(t) \\ - \ & \mathcal{V}(y)(t)]\int_t^{t_n}\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(\nu)d\nu\}dt.\end{aligned}$$

So we can arrive at

$$\begin{aligned}||\wp_J(\tilde{u}_n^h - \bar{u}_n^h)(t)||_{0,n} &\leq h_n^{1/2}|\tilde{U}_{n-1} - y(t_{n-1})| + h_n^{1/2}||a(t)||_{0,n}||\tilde{u}_n^h - \bar{u}_n^h||_{0,n} \\ &+ [1 + h_n^{1/2}||a(t)||_{0,n}]||\bar{u}_n^h - y||_{0,n} \\ &+ Lh_n^{1/2}\int_{I_n}|\tilde{\mathcal{V}}(\tilde{u}^h)(t) - \mathcal{V}(y)(t)|dt. \qquad (2.6.2.12)\end{aligned}$$

We again use Lemma 2.6.1.1, where we have now replaced u by $\tilde{u}$. Recalling (2.6.2.12), we obtain

$$\begin{aligned}
& [\beta_1 - h_n^{1/2}||a(t)||_{0,n} - Ch_n^{3/2}]||\tilde{u}_n^h - \bar{u}_n^h||_{0,n} \\
\leq\ & h_n^{1/2}|\tilde{U}_{n-1}^h - y(t_{n-1})| + [1 + h_n^{1/2}||a(t)||_{0,n}]||\bar{u}_n^h - y||_{0,n} \\
+\ & Ch_n^{3/2}||\bar{u}_n^h(t) - y(t)||_{0,n} + h_n^{1/2}\sum_{l=1}^{J}|\tilde{u}_n^h(t_{n_l}) - \bar{u}_n^h(t_{n_l})| \\
+\ & Ch_n^{3/2}\{|[\tilde{\mathcal{V}}(\tilde{u}^h)(t_{n-1}) - \tilde{\mathcal{V}}(y)(t_{n-1})]| + |[\tilde{\mathcal{V}}(y)(t_n) - \mathcal{V}(y)(t_n)]|\}.
\end{aligned}$$

Hence

$$\begin{aligned}
||\tilde{u}_n^h - y||_{0,n} \ \leq\ & ||\tilde{u}_n^h - \bar{u}_n^h||_{0,n} + ||\bar{u}_n^h - y||_{0,n} \\
\leq\ & Ch_n^{1/2}\{|\tilde{U}_{n-1}^h - y(t_{n-1})| + \sum_{l=1}^{J}|\tilde{u}_n^h(t_{n_l}) - \bar{u}_n^h(t_{n_l})|\} \\
+\ & C||\bar{u}_n^h(t) - y(t)||_{0,n} + Ch_n^{3/2}\sum_{\ell=1}^{n-1}|\tilde{U}_\ell^h - y(t_\ell)| \\
+\ & Ch_n^{3/2}|\tilde{\mathcal{V}}(y)(t_n) - \mathcal{V}(y)(t_n)|. \qquad (2.6.2.13)
\end{aligned}$$

Substitute $\tilde{v}_n^h = \tilde{U}_n^h - y(t_n)$ into (2.6.2.11)

$$\begin{aligned}
|\tilde{U}_n^h - y(t_n)| \ \leq\ & |\tilde{U}_{n-1}^h - y(t_{n-1})| + ||a(t)||_{0,n}||\tilde{u}_n^h - y||_{0,n} \\
+\ & Ch_n^{3/2}||\tilde{u}_n^h(t) - y(t)||_{0,n} + Ch_n^{3/2}\sum_{\ell=1}^{n-1}|\tilde{U}_\ell^h - y(t_\ell)| \\
+\ & Ch_n^{3/2}|\tilde{\mathcal{V}}(y)(t_n) - \mathcal{V}(y)(t_n)|. \qquad (2.6.2.14)
\end{aligned}$$

We therefore obtain

$$|\tilde{U}_n^h - y(t_n)| \ \leq\ Ch_n^{1/2}||\tilde{u}_n^h - y||_{0,n} + Ch_n^{3/2}|\tilde{\mathcal{V}}(y)(t_n) - \mathcal{V}(y)(t_n)|. \qquad (2.6.2.15)$$

We note that for $J \leq M$ and the $u_n^h(t_{n_l}) = U_{n_l}^h$, inequality (2.6.2.13) can be rearranged in the form

$$\begin{aligned}
||\tilde{u}_n^h - y||_{0,n} &\leq Ch_n^{1/2}\sum_{\ell=1}^{n-1}|\tilde{U}_\ell^h - y(t_\ell)| + Ch^{m+1}||y^{(m+1)}||_{0,\bar{n}} \\
&+ Ch_n^{3/2}|\tilde{\mathcal{V}}(y)(t_n) - \mathcal{V}(y)(t_n)| \\
&\leq Ch_n^{1/2}\sum_{\ell=1}^{n-1}h_\ell^{1/2}||\tilde{u}_\ell^h - y||_{0,\ell} + C\sum_{\ell=1}^{n}h_\ell^{3/2}|\tilde{\mathcal{V}}(y)(t_\ell) - \mathcal{V}(y)(t_\ell)| \\
&+ Ch^{m+1}||y^{(m+1)}||_{0,\bar{n}}. \qquad (2.6.2.16)
\end{aligned}$$

Therefore,

$$||\tilde{u}_n^h - y||_{0,n} \leq Ch^{m+1}||y^{(m+1)}||_{0,\bar{n}} + C\sum_{l=1}^{n}h_l^{3/2}|\tilde{\mathcal{V}}(y)(t_l) - \mathcal{V}(y)(t_l)|.$$

Here $||\cdot||_{0,\bar{n}}$ is the L^2-norm over $[t_{n-M}, t_n]$. Thus (2.6.2.8) and (2.6.2.10) hold true when $j = 0$. From (2.6.2.14) and Gronwall's lemma we also know that (2.6.2.7) and (2.6.2.9) hold. Inequalities (2.6.2.10) for $1 \leq j \leq m+1$ are obtained by using the inequalities

$$||\tilde{u}_n^h - y||_{j,n} \leq ||\tilde{u}_n^h - \bar{u}_n^h||_{j,n} + ||\bar{u}_n^h - y||_{j,n}$$

and

$$||\tilde{u}_n^h - \bar{u}_n^h||_{j,n} \leq Ch^{-j}||\tilde{u}_n^h - \bar{u}_n^h||_{0,n}.$$

The following theorem is the analogue of Theorem 2.6.1.2.

Theorem 2.6.2.2. *Assume that the assumptions of Theorem 2.6.2.1 hold and take $\{t_n\}_{n=1}^N$ as the Gaussian points in $I := [0,T]$. Then for sufficiently small h we have*

$$\max\{|\tilde{U}_n^h - y(t_n)| : \ n = 0, \ldots, N\} \leq Ch^{2m+2-J}.$$

Proof. Choose $\tilde{\bar{v}}^h = (\tilde{V}_0, \tilde{v}_1^h, \ldots, \tilde{v}_N^h)$ such that $\tilde{v}_1^h(t_0) = \tilde{V}_0$ and $\tilde{v}_n^h(t_n) = \tilde{v}_{n+1}^h(t_n)$, $n = 1, \ldots, N-1$. Substitute that $\tilde{\bar{v}}^h$ in (2.6.2.11) and sum up over $j = 1, \ldots, n$. This yields

$$(\tilde{U}_n^h - y(t_n))\tilde{v}_n^h(t_n) = \int_0^{t_n} (\tilde{u}^h - y)(\tilde{v}^h)' dt - \int_0^{t_n} a(t)(\tilde{u}^h - y)\tilde{v}^h dt + \int_0^{t_n} [\tilde{\mathcal{V}}(\tilde{u}^h) - \mathcal{V}(y)]\tilde{v}^h dt. \qquad (2.6.2.17)$$

Let $\tilde{w}$ in $H^{m+2-J}([0,t];\Omega)$ be the solution of

$$\begin{cases} \tilde{w}' - [a(t) + \omega_{nn} k(0)t]\tilde{w} = 0, \\ \tilde{w}(t_n) = \tilde{U}_n^h - y(t_n), \end{cases}$$

for $t \in [0, t_n]$. Let $\tilde{w}^h$ be a continuous piecewise interpolating polynomial of degree $m + 1 - J$ of $\tilde{w}$ such that $\tilde{w}^h(t_n) = \tilde{U}_n^h - y(t_n)$. From Sobolev interpolation theory [22], we deduce

$$||\tilde{w} - \tilde{w}^h||_1 \leq Ch^{m+1-J}||\tilde{w}^{(m+2-J)}||_0.$$

Now we use the following Lemma 2.6.2.1 to express the norm of $\tilde{w}^{(m+2-J)}$ in terms of $\tilde{U}_n^h - y(t_n)$.

Lemma 2.6.2.1. *Fix s in $[0, T]$. Suppose that g is such that $g|_{I_n} \in H^{m+1-J}(I_n;\Omega)$ for $n = 1, \ldots, N$, and that $\tilde{w}$ in $H^1([0,T];\Omega)$ satisfies the equation*

$$\tilde{w}'(t) - [a(t) + \omega_{nn} k(0)t]\tilde{w}(t) = g(t),$$

for $t \in [0, t_n]$. Then there exists a constant C independent of g and $s \in [0, t_n]$ such that

$$||\tilde{w}||_{m+2-J} \leq C\{|\tilde{w}(s)| + ||g||_{m+1-J}\}.$$

Proof. The lemma is a special case of Lemma 2.6.1.2 without the memory term. Compare also with Delfour and Dubeau [40].

We then have

$$||\tilde{w} - \tilde{w}^h||_1 \leq Ch^{m+1-J}|\tilde{U}_n^h - y(t_n)|. \qquad (2.6.2.18)$$

Set $\tilde{v}^h = \tilde{w}^h$ in (2.6.2.17)

$$\begin{aligned} |\tilde{U}_n^h - y(t_n)|^2 &\leq \int_0^{t_n} (\tilde{u}^h - y)[(\tilde{w} - \tilde{w}^h)' - a(t)(\tilde{w} - \tilde{w}^h)]dt \\ &- \int_0^{t_n} \tilde{w}\omega_{nn}t(\tilde{u}^h - y)dt + \int_0^{t_n} [\tilde{\mathcal{V}}(\tilde{u}^h) - \mathcal{V}(y)]\tilde{w}^h dt. \end{aligned} \tag{2.6.2.19}$$

Since

$$\begin{aligned} \tilde{\mathcal{V}}(\tilde{u}^h) - \mathcal{V}(y) &= \tilde{\mathcal{V}}(\tilde{u}^h) - \tilde{\mathcal{V}}(y) + \tilde{\mathcal{V}}(y) - \mathcal{V}(y) \\ &= \tilde{\mathcal{V}}(y) - \mathcal{V}(y) + \sum_{i=1}^{n-1} \omega_{ni}(\tilde{U}_i^h - y(t_i)) \\ &+ \omega_{nn}t(\tilde{u}^h(t) - y(t)), \end{aligned} \tag{2.6.2.20}$$

and from Lemma 2.6.2.1 we have

$$\begin{aligned} ||\tilde{w}^h||_0 &\leq ||\tilde{w} - \tilde{w}^h||_0 + ||\tilde{w}||_0 \\ &\leq Ch^{m+2-J}|\tilde{U}_n^h - y(t_n)| + C|\tilde{U}_n^h - y(t_n)|, \end{aligned} \tag{2.6.2.21}$$

Combining {(2.6.2.18), (2.6.2.19), (2.6.2.20), (2.6.2.21), and Lemma 2.6.2.1} yields

$$\begin{aligned} |\tilde{U}_n^h - y(t_n)| &\leq Ch^{m+1-J}||\tilde{u}^h - y||_0 + Ch^{2m+2-J}||y^{(m)}||_0 \\ &+ C\sum_{i=1}^{n-1} \omega_{ni}k(t_{n-1} - t_i)|\tilde{U}_i^h - y(t_i)|. \end{aligned} \tag{2.6.2.22}$$

Hence, Gronwall's lemma leads to

$$|\tilde{U}_n^h - y(t_n)| \leq Ch^{2m+2-J}.$$

Remark 2.6.2.1. *We conclude that it is not substantially more difficult to analyze the mesh-dependent Galerkin method for nonstandard Volterra integro-differential equations containing memory terms of the form* $\mathcal{V}_G^N(y)(t) := \int_0^t k(t-s)G(y(t), y(s))ds$. *We leave the details to the interested readers.*

Chapter 3

The discontinuous Galerkin method for delay VIDEs

In this chapter we focus on three kinds of delay Volterra integro-differential equations. We show the regularities of those problems and thus construct and analyze the robust adaptive discontinuous Galerkin methods for them. The readers may wish to consult Brunner and Zhang [28], Hale [58], Hale and Verduyn Lunel [59], and Bellen and Zennaro [12] and the references therein for the background materials and related results about delay differential or integro-differential equations.

3.1 Primary discontinuities of several classes of delay Volterra integro-differential equations

3.1.1 Delay VIDEs with weakly singular kernels

Let us consider

$$\begin{cases} y'(t) = f(t, y(t)) + \int_0^t (t-s)^{-\alpha} G(s, y(s), y(\theta(s)))ds, \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.1.1}$$

where we assume $0 < \alpha < 1$ and

(i) f, G, ϕ are sufficiently smooth.

(ii) $\theta(t) := t - \tau(t)$, with τ sufficiently smooth, $t > \tau(t) \geq \tau_0 > 0$ $(t \in I)$. Moreover, θ is strictly increasing on I and $\bar{a} = \inf_{t\geq 0} \theta(t) < 0$.

(iii) The points $\{\xi_\mu\}$ are defined by

$$\theta(\xi_\mu) = \xi_\mu - \tau(\xi_\mu) =: \xi_{\mu-1}, \quad \forall \mu \geq 1, \tag{3.1.1.2}$$

where $\xi_0 := 0$. Obviously

$$\xi_{\mu+1} - \xi_\mu \geq \tau_0 > 0, \quad \forall \mu \geq 0.$$

For simplicity we denote $\overline{G}(s) := G(s, y(s), y(\theta(s)))$. We shall use the following formula frequently.

$$\begin{aligned} H(t) &:= \int_\xi^t (t-s)^{-\alpha} \overline{G}(s) ds \\ &= \frac{1}{1-\alpha} \overline{G}(\xi)(t-\xi)^{1-\alpha} + \frac{1}{(1-\alpha)(2-\alpha)} \overline{G}'(\xi)(t-\xi)^{2-\alpha} \\ &+ \cdots + \frac{1}{(1-\alpha)_{m+1}} \overline{G}^{(m)}(\xi)(t-\xi)^{m+1-\alpha} \\ &+ \frac{1}{(1-\alpha)_{m+1}} \int_0^t \overline{G}^{(m+1)}(s)(t-s)^{m+1-\alpha} ds, \end{aligned} \tag{3.1.1.3}$$

with $(1-\alpha)_m := (1-\alpha)(2-\alpha)\cdots(m-\alpha)$. We remark that (3.1.1.3) can be obtained by using repeated integration by parts.

Definition 3.1.1.1. If the solution of (3.1.1.1) and its derivatives of order less than, or equal to μ are continuous at some points $\xi \in I$ but the derivatives of order $\mu + 1$ is not, then ξ is called *a primary discontinuity* of problem (3.1.1.1)

Denote $J^{[\mu]} := (\xi_{\mu-1}, \xi_{\mu+1}]$ $(\mu \geq 0)$, where $\xi_{-1} := \bar{a}$. We shall describe the primary discontinuities in solutions for (3.1.1.1) as the following theorem.

Theorem 3.1.1.1. *The primary discontinuities of problem (3.1.1.1) are the points* ξ_μ $(\mu = 0, 1, \ldots)$ *generated by (3.1.1.2). To be more precise,* $y \in C^{2\mu,1-\alpha}(J^{[\mu]})$*, but* $y^{(2\mu+1)}$ *is not continuous at the point* ξ_μ*, provided the assumptions* (i) *and* (ii) *hold.*

Remark 3.1.1.1. *We use* $C^\beta(I)$ $(0 < \beta < 1)$ *to denote the well-known Hölder space:* V *is in* $C^\beta(I)$ *if, for any* $t_1,\ t_2 \in I$ $(t_1 \neq t_2)$*, we have*

$$|V(t_1) - V(t_2)| \leq L \cdot |t_1 - t_2|^\beta.$$

A function V *is in* $C^{\mu,\beta}(I)$ $(\mu \in \mathbb{N}, \mu \geq 1)$ *if* $V \in C^\mu(I)$ *and* $V^{(\mu)} \in C^\beta(I)$*. We set* $C^{0,\beta}(I) = C^\beta(I)$*.*

Proof. The proof is based on the method of steps.

(1) Consider the regularity of the solution for (3.1.1.1) at the point $\xi_0 := 0$. It is possible to satisfy the condition $y(0) = \phi(0)$, but not, in general, also the condition $y'(0+) = \phi'(0-)$. The continuity of the derivative of the solution can be guaranteed at the initial point 0 only for deliberately chosen $\phi(t)$, and such a function $\phi(t)$ must satisfy the condition $\phi'(0-) = f(0, \phi(0))$.

(2) Consider the regularity at the point ξ_1. We write the equation (3.1.1.1) as

$$\begin{cases} y'(t) = f(t, y(t)) + H_1(t), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases}$$

where

$$H_1(t) := \int_0^t (t-s)^{-\alpha} \overline{G}(s) ds.$$

From the formula (3.1.1.3), we obtain

$$H_1(t) = \frac{1}{1-\alpha} \overline{G}(0) t^{1-\alpha} + \frac{1}{1-\alpha} \int_0^t \overline{G}'(s)(t-s)^{1-\alpha} ds.$$

Thus

$$H_1'(t) = \overline{G}(0) t^{-\alpha} + \int_0^t \overline{G}'(s)(t-s)^{-\alpha} ds.$$

So y'' is continuous at the points ξ_μ $(\mu \geq 1)$. But $y^{(3)}$ is discontinuous at ξ_1. Now we are going to prove that

$$\int_0^t \overline{G}'(s)(t-s)^{-\alpha}ds \in C^{1-\alpha}(J^{[1]}). \tag{3.1.1.4}$$

For any $t_0,\ t_1 \in J^{[1]}$ and without loss of generality we assume $t_0 \leq \xi_1 \leq t_1$. From

$$\begin{aligned}
& \left|\int_0^{t_1} \overline{G}'(s)(t_1-s)^{-\alpha}ds - \int_0^{t_0} \overline{G}'(s)(t_0-s)^{-\alpha}ds\right| \\
\leq & \left|\int_{t_0}^{t_1} \overline{G}'(s)(t_1-s)^{-\alpha}ds\right| + \left|\int_0^{t_0} \left(\overline{G}'(s)(t_1-s)^{-\alpha} - \overline{G}'(s)(t_0-s)^{-\alpha}\right) ds\right| \\
\leq & L_{\overline{G}'} \left\{\frac{2(t_1-t_0)^{1-\alpha}}{1-\alpha} + \frac{t_0^{1-\alpha}}{1-\alpha} - \frac{t_1^{1-\alpha}}{1-\alpha}\right\} \\
\leq & \frac{3L_{\overline{G}'}}{1-\alpha}(t_1-t_0)^{1-\alpha},
\end{aligned}$$

where $L_{\overline{G}'}$ is the upper bound of $|\overline{G}'(s)|$ in $J^{[1]}$. So $y \in C^{3-\alpha}(J^{[1]})$.

(3) Consider now the regularity at the point ξ_2. We write the equation (3.1.1.1) as

$$\begin{cases} y'(t) = f(t, y(t)) + \displaystyle\int_0^{\xi_1} (t-s)^{-\alpha}\overline{G}(s)ds + H_2(t), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases}$$

where

$$H_2(t) := \int_{\xi_1}^t (t-s)^{-\alpha}\overline{G}(s)ds.$$

By using the formula (3.1.1.3) again, we obtain

$$\begin{aligned}
H_2(t) \ = \ & \frac{1}{1-\alpha}\overline{G}(\xi_1)(t-\xi_1)^{1-\alpha} + \frac{1}{(1-\alpha)(2-\alpha)}\overline{G}'(\xi_1)(t-\xi_1)^{2-\alpha} \\
+ \ & \frac{1}{(1-\alpha)(2-\alpha)(3-\alpha)}\overline{G}''(\xi_1)(t-\xi_1)^{3-\alpha} \\
+ \ & \frac{1}{(1-\alpha)(2-\alpha)(3-\alpha)}\int_{\xi_1}^t \overline{G}^{(3)}(t-s)^{3-\alpha}ds.
\end{aligned}$$

Thus we can calculate the derivatives of order up to three of $H_2(t)$ as:

$$\begin{aligned}
H_2'(t) &= \overline{G}(\xi_2 1)(t-\xi_1)^{-\alpha} + \frac{1}{1-\alpha}\overline{G}'(\xi_1)(t-\xi_1)^{1-\alpha} \\
&+ \frac{1}{(1-\alpha)(2-\alpha)}\overline{G}''(\xi_1)(t-\xi_1)^{2-\alpha} \\
&+ \frac{1}{(1-\alpha)(2-\alpha)}\int_{\xi_1}^{t}\overline{G}^{(3)}(s)(t-s)^{2-\alpha}ds. \\
H_2''(t) &= -\alpha\overline{G}(\xi_1)(t-\xi_1)^{-\alpha-1} + \overline{G}'(\xi_1)t(t-\xi_1)^{-\alpha} + \frac{1}{1-\alpha}\overline{G}''(\xi_1)(t-\xi_1)^{1-\alpha} \\
&+ \frac{1}{1-\alpha}\int_{\xi_1}^{t}\overline{G}^{(3)}(s)(t-s)^{(1-\alpha)}ds. \\
H_2^{(3)}(t) &= \alpha(\alpha+1)\overline{G}(\xi_1)(t-\xi_1)^{-\alpha-2} - \alpha\overline{G}'(\xi_1)(t-\xi_1)^{-\alpha-1} \\
&+ \overline{G}''(\xi_1)(t-\xi_1)^{-\alpha} + \int_{\xi_1}^{t}\overline{G}^{(3)}(s)(t-s)^{-\alpha}ds.
\end{aligned}$$

Hence $y^{(4)}$ is continuous at the point ξ_μ ($\mu \geq 2$). But $y^{(5)}$ is discontinuous at ξ_2. Furthermore along the lines proving (3.1.1.4) we can verify that

$$\int_{\xi_1}^{t}\overline{G}^{(3)}(t-s)^{-\alpha}ds \in C^{1-\alpha}(J^{[2]}).$$

(4) We suppose $y \in C^{2\mu,1-\alpha}(J^{[\mu]})$ and $y \in C^{2\mu}(\xi_m)$ ($m \geq \mu$). Now we consider the regularity at $\xi_{\mu+1}$. We write the equation (3.1.1.1) as

$$\begin{cases} y'(t) = f(t,y(t)) + \int_0^{\xi_\mu}(t-s)^{-\alpha}\overline{G}(s)ds + H_{\mu+1}(t), \quad t \in I := [0,T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases}$$

where

$$H_{\mu+1}(t) := \int_{\xi_\mu}^{t}(t-s)^{-\alpha}\overline{G}(s)ds.$$

We write $H_{\mu+1}(t)$ as, by using (3.1.1.3),

$$\begin{aligned}
H_{\mu+1}(t) &= \frac{1}{1-\alpha}\overline{G}(\xi_\mu)(t-\xi_\mu)^{1-\alpha} + \frac{1}{(1-\alpha)(2-\alpha)}\overline{G}'(\xi_\mu)(t-\xi_\mu)^{2-\alpha} \\
&+ \cdots + \frac{1}{(1-\alpha)_{2\mu+1}}\overline{G}^{(2\mu)}(\xi_\mu)(t-\xi_\mu)^{2\mu+1-\alpha} \\
&+ \frac{1}{(1-\alpha)_{2\mu+1}}\int_{\xi_\mu}^{t}\overline{G}^{(2\mu+1)}(s)(t-s)^{2\mu+1-\alpha}ds. \qquad (3.1.1.5)
\end{aligned}$$

Thus we can calculate the derivatives of order up to $2\mu+1$ of $H_{\mu+1}(t)$ as:

$$\begin{aligned} H'_{\mu+1}(t) &= \overline{G}(\xi_\mu)(t-\xi_\mu)^{-\alpha} + \frac{1}{1-\alpha}\overline{G}'(\xi_\mu)(t-\xi_\mu)^{1-\alpha} \\ &+ \cdots + \frac{1}{(1-\alpha)_{2\mu}}\overline{G}^{(2\mu)}(\xi_\mu)(t-\xi_\mu)^{2\mu-\alpha} \\ &+ \frac{1}{(1-\alpha)_{2\mu}}\int_{\xi_\mu}^{t}\overline{G}^{(2\mu+1)}(s)(t-s)^{2\mu-\alpha}ds. \end{aligned} \tag{3.1.1.6}$$

$$\vdots$$

$$\begin{aligned} H^{(2\mu+1)}_{\mu+1}(t) &= (\alpha)_{2\mu}\overline{G}(\xi_\mu)(t-\xi_\mu)^{-\alpha-2\mu} + \cdots + \overline{G}^{(2\mu)}(\xi_\mu)(t-\xi_\mu)^{-\alpha} \\ &+ \frac{1}{(1-\alpha)_{2\mu}}\int_{\xi_\mu}^{t}\overline{G}^{(2\mu+1)}(s)(t-s)^{-\alpha}ds. \end{aligned} \tag{3.1.1.7}$$

Hence $y \in C^{2(\mu+1),1-\alpha}(J^{[\mu+1]})$.

Let us consider now

$$\begin{cases} y'(t) = f(t, y(t), y(\theta(t))) + \int_0^t (t-s)^{-\alpha} G(s, y(s), y(\theta(s)))ds, & t \in I := [0,T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.1.8}$$

with the assumptions (i), (ii), (iii) in (3.1.1.1).

Theorem 3.1.1.2. *The primary discontinuities of problem (3.1.1.8) are the points* ξ_μ ($\mu = 0, 1, \ldots$) *generated by (3.1.1.2). To be more precise,* $y \in C^{\mu,1-\alpha}(J^{[\mu]})$, *but* $y^{(\mu+1)}$ *is not continuous at the point* ξ_μ, *provided the assumptions* (i) *and* (ii) *hold.*

Proof. From [125] we know the primary discontinuities of problem

$$\begin{cases} y'(t) = f(t, y(t), y(\theta(t))), & t \in I := [0,T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.1.9}$$

are the points ξ_μ ($\mu = 0, 1, \ldots$) generated by (3.1.1.2). To be more precise, $y^{(\mu)}$ is continuous at the point ξ_μ and $y^{(\mu+1)}(\xi_{\mu+1})$ is bounded, but $y^{(\mu+1)}$ is not continuous at ξ_μ, provided the assumptions (i) and (ii) hold. Thus, recalling Theorem 3.1.1.1, we obtain Theorem 3.1.1.2.

Theorem 3.1.1.3. *Consider*

$$\begin{cases} y'(t) = f(t, y(t), y(\theta(t)), y'(\theta(t))) + \displaystyle\int_0^t (t-s)^{-\alpha} G(s, y(s), y(\theta(s)))ds, & t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0]. \end{cases} \tag{3.1.1.10}$$

with the assumptions (i), (ii), (iii) *in (3.1.1.1). There is no smoothing to the solutions of (3.1.1.10); more precisely,* $y \in C^{1-\alpha}(J^{[\mu]}),\ \forall \mu \geq 1$.

Proof. It is known that there is no smoothing to the solutions of neutral delay differential equations (see [125]), so there is no smoothing to the solutions of (3.1.1.10). Following the lines in the proof of Theorem 3.1.1.1, we can verify that

$$y \in C^{1-\alpha}(J^{[\mu]}),\ \forall \mu \geq 1.$$

Remark 3.1.1.2. *The delay integro-differential equations may include terms such as*

$$\int_0^t K(t-s)G(y(\theta(s)))ds, \tag{3.1.1.11}$$

and

$$\int_0^{\theta(t)} K(t-s)G(y(s))ds. \tag{3.1.1.12}$$

A natural question arises: "What is the difference between the regularity of the delay integro-differential equation with the term (3.1.1.11) and that with the term (3.1.1.12)?".

The term (3.1.1.11) can be expressed as

$$\int_{\theta(0)}^{\theta(t)} K(t-\theta^{-1}(s))(\theta^{-1}(s))'G(y(s))ds. \tag{3.1.1.13}$$

Therefore, if the kernel function K is sufficiently smooth, then the regularities of (3.1.1.11) and (3.1.1.12) are the same. But when K has weakly singular behavior, their regularities will be different. Discussion of Section 3.1.2 will explain this well.

3.1.2 Delay functional VIDEs of Hale's type

Consider

$$\begin{cases} \dfrac{d}{dt}\left(y(t) - \displaystyle\int_0^{\theta(t)} K(t-s)G(y(s))ds\right) = f(t, y(t), y(\theta(t))), & t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.2.1}$$

where we assume that

(a) K is sufficiently smooth, and

(b) (i), (ii), (iii) in (3.1.1.1) hold.

We describe the primary discontinuities of problem (3.1.2.1) in the following theorem.

Theorem 3.1.2.1. *The primary discontinuities of problem (3.1.2.1) are the points ξ_μ ($\mu = 0, 1, \ldots$) generated by (3.1.1.2). To be more precise, $y^{(\mu)}$ is continuous at ξ_μ but $y^{(\mu+1)}$ is, in general, not provided the assumptions* (i) *and* (ii) *hold.*

Proof. We rewrite the left-hand side of (3.1.2.1) as

$$\begin{aligned} \frac{d}{dt}\left(y(t) - \int_0^{\theta(t)} K(t-s)G(y(s))ds\right) &= y'(t) - \theta'(t)K(t-\theta(t))G(y(\theta(t))) \\ &- \int_0^{\theta(t)} K_t(t-s)G(y(s))ds. \end{aligned} \tag{3.1.2.2}$$

Thus the remaining lines of the proof are easily generated by using the method of steps (cf. Brunner and Zhang [28]).

For the equation:

$$\begin{cases} \dfrac{d}{dt}\left(y(t) - \displaystyle\int_0^{\theta(t)} (t-s)^{-\alpha}G(y(s))ds\right) = f(t, y(t), y(\theta(t))), & t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.2.3}$$

where $1 < \alpha < 1$ and assuming (i), (ii), (iii) in (3.1.2.1) hold, we have Theorem 3.1.2.2.

Theorem 3.1.2.2. *The primary discontinuities of problem (3.1.2.3) are the points ξ_μ $(\mu = 0, 1, \ldots)$ generated by (3.1.1.2). To be more precise, $y^{(\mu)}$ is continuous at ξ_μ but $y^{(\mu+1)}$ is, in general, not provided the assumptions* (i) *and* (ii) *hold. If, in addition, we assume $y \in C^{1-\alpha}(J^{[0]})$, then $y \in C^{\mu,1-\alpha}(J^{[\mu]})$, but $y^{(\mu+1)}$ is not continuous at $\xi_{\mu+1}$.*

Proof. Since

$$\frac{d}{dt}\left(y(t) - \int_0^{\theta(t)} (t-s)^{-\alpha} G(y(s))ds\right) \tag{3.1.2.4}$$

$$= \quad y'(t) - \theta'(t)(t-\theta(t))^{-\alpha}G(y(\theta(t))) + \alpha \int_0^{\theta(t)} (t-s)^{-1-\alpha}G(y(s))ds,$$

and $t - \theta(t) = \tau(t) \geq \tau_0 > 0$ hold, assuming that y is continuous at the point $\xi_0 := 0$, we can realize the assertion with the method of steps as in Brunner and Zhang [28]. If we assume, in addition, $y \in C^{1-\alpha}(J^{[0]})$, then we can prove the second assertion still by using the method of steps.

We now consider

$$\begin{cases} \dfrac{d}{dt}\left(y(t) - \displaystyle\int_{\theta(t)}^{t} K(t-s)G(y(s))ds\right) = f(t, y(t), y(\theta(t))), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.2.5}$$

where we assume K is sufficiently smooth and (i), (ii), (iii) in (3.1.1.1) hold. The primary discontinuities of problem (3.1.2.5) are described by Theorem 3.1.2.3.

Theorem 3.1.2.3. *The primary discontinuities of problem (3.1.2.5) are the points ξ_μ $(\mu = 0, 1, \ldots)$ generated by (3.1.1.2). More precisely, $y^{(\mu)}$ is continuous at ξ_μ but $y^{(\mu+1)}$ is not, in general, provided the assumptions* (i) *and* (ii) *hold.*

Proof. We rewrite the left-hand side of (3.1.2.5)

$$\begin{aligned}
&\frac{d}{dt}\left(y(t) - \int_{\theta(t)}^{t} K(t-s)G(y(s))ds\right) \\
= \;& y'(t) + \theta'(t)K(t-\theta(t))G(y(\theta(t))) - K(0)G(y(t)) \\
- \;& \int_{\theta(t)}^{t} K_t(t-s)G(y(s))ds. \qquad (3.1.2.6)
\end{aligned}$$

Hence the proof can be completed by using the method of steps (cf. Brunner and Zhang [28]).

Consider now

$$\begin{cases} \dfrac{d}{dt}\left(y(t) - \displaystyle\int_{\theta(t)}^{t} (t-s)^{-\alpha}G(y(s))ds\right) = f(t, y(t), y(\theta(t))), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \qquad (3.1.2.7)$$

where we assume $0 < \alpha < 1$ and (i), (ii), (iii) in (3.1.1.1). This case is different from the one treated in Theorem 3.1.2.2, since we cannot use the techniques in the proof of Theorem 3.1.2.1. We describe the primary discontinuities of problem (3.1.2.7) in the following theorem.

Theorem 3.1.2.4. *There is no smoothing to the solution of (3.1.2.7). To be more precise.* $y \in C^{1,1-\alpha}(J^{[\mu]})$, *for all* $\mu \geq 1$, *where* C *is independent on* μ, *provided the assumptions* (i) *and* (ii) *hold.*

Proof. Consider first the regularity of the solution for (3.1.2.7) at the point $\xi_0 := 0$. It is possible to choose $y(0) = \phi(0)$. The continuity of the derivative of the solution can be guaranteed at the initial point $\xi_0 := 0$ only for $\phi(t)$ satisfying the condition

$$\phi'(0-) = \frac{d}{dt}\left(\int_{\theta(t)}^{t} (t-s)^{-\alpha}G(\phi(s))ds\right) + f(0, \phi(0), \phi(\theta(0))).$$

Consider now the regularity at the point ξ_1. We write the equation (3.1.2.7) as

$$\begin{cases} y'(t) = f(t, y(t), y(\theta(t))) + \dfrac{d}{dt}\left(H_1(t)\right), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases}$$

where

$$H_1(t) := \int_{\theta(t)}^{t} (t-\theta(t))^{-\alpha}\overline{G}(s)ds,$$

where $\overline{G}(s) := G(y(s))$. From the formula (3.1.1.3), we obtain

$$H_1(t) = \frac{1}{1-\alpha}\overline{G}(\theta(t))t^{1-\alpha} + \frac{1}{1-\alpha}\int_{\theta(t)}^{t} \overline{G}'(s)(t-s)^{1-\alpha}ds.$$

Thus

$$\frac{d}{dt}(H_1(t)) = \overline{G}(\theta(t))(t-\theta(t))^{-\alpha}(1-\theta'(t)) + \int_{\theta(t)}^{t} \overline{G}'(s)(t-s)^{-\alpha}ds,$$

where $\overline{G}'(s) = G_y(y(s))y'(s)$. Similarly, we calculate that

$$\begin{aligned}\frac{d^2(H_1(t))}{dt^2} &= [\overline{G}(\theta(t))\cdot(t-\theta(t))^{-\alpha}\cdot(1-\theta'(t))]' + \overline{G}'(\theta(t))(t-\theta(t))^{-\alpha}(1-\theta'(t))\\ &+ \int_{\theta(t)}^{t} \overline{G}''(s)(t-s)^{-\alpha}ds.\end{aligned}$$

Since

$$\overline{G}'(\theta(t)) = G_y(y(\theta(t)))y'(\theta(t)).$$

we see that y'' is discontinuous at the point ξ_1 and thus there is no smoothing to the solution of (3.1.2.7). To prove that $y \in C^{1,1-\alpha}(J^{[\mu]})$, it is sufficient to verify that

$$\left|\int_{\theta(t)}^{t} (t-s)^{-\alpha}\overline{G}'(s)ds\right| \le C(t-\xi_1)^{1-\alpha}. \tag{3.1.2.8}$$

Since $\overline{G}'(t)$ is continuous in $[\bar{a}, T]$, we have

$$\begin{aligned}\left|\int_{\theta(t)}^{t} (t-s)^{-\alpha}\overline{G}'(s)ds\right| &\le L_G\int_{\theta(t)}^{t} (t-s)^{-\alpha}ds\\ &\le C(t-\theta(t))^{1-\alpha}\\ &\le C(t-\xi_1)^{1-\alpha}. \end{aligned} \tag{3.1.2.9}$$

In the last step of (3.1.2.9), we use that θ is strictly increasing in I and thus

$$\xi_1 = \theta(\xi_2) > \theta(t) > \theta(\xi_1) = \xi_0, \quad \forall t \in (\xi_1, \xi_2).$$

Following the above process we can establish that

$$y'(t) \leq C(t - \xi_\mu)^{1-\alpha}, \ \forall \mu \geq 1.$$

3.1.3 Delay VIDEs of neutral type with smooth kernels

Consider the following delay VIDEs of neutral type:

$$\begin{cases} y'(t) = f(t, y(t), y(\theta(t))) + \displaystyle\int_0^t K(t-s)G(s, y(s), y'(s), y'(\theta(s)))ds, \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.3.1}$$

where we assume that

(a) K is sufficiently smooth, and

(b) (i), (ii), (iii) in (3.1.1.1) hold.

We describe the primary discontinuities of problem (3.1.3.1) in the following theorem.

Theorem 3.1.3.1. *The primary discontinuities of problem (3.1.3.1) are generated inductively by the recursion (3.1.1.2), where $\xi_0 := 0$. More precisely, $y^{(\mu)}$ and lower-order derivatives are continuous at ξ_μ, but $y^{(\mu+1)}$ is ,in general, not under the assumptions (a) and (b) for (3.1.3.1).*

Proof. (1) We consider the regularity of the solution for (3.1.3.1) at the point $\xi_0 := 0$. It is possible to choose ϕ to satisfy $y(0) = \phi(0)$. However in general, $\phi'(0-) \neq y'(0+)$. Hence y is continuous at ξ_μ ($\mu \geq 0$), but y' is not continuous at ξ_0.

(2) Consider the regularity at ξ_1. Obviously, y' is continuous at ξ_μ ($\mu \geq 1$). But since the epression for y'' at the point ξ_1 include y' at the point ξ_0, y'' is not continuous at the point ξ_1.

(3) Consider the regularity at ξ_2. Set $\overline{G}(s) := G(s, y(s), y'(s), y'(\theta(s)))$. We obtain

$$y''(t) = f'(t, y(t), y(\theta(t))) + K(0)\overline{G}(t) + \int_0^t K_t(t-s)\overline{G}(s)ds.$$

Hence y'' is continuous at ξ_μ $(\mu \geq 2)$. But $y^{(3)}$ is not continuous at ξ_2, since $y^{(3)}$ includes $y^{(2)}$ which is not continuous at ξ_2.

(4) Suppose $y^{(\mu)}$ is continuous at ξ_m $(m \geq \mu)$. Consider now the regularity at $\xi_{\mu+1}$. We know that $y^{(\mu+1)}(\xi_\mu)$ includes $y^{(m)}(\xi_\mu)$ $(m \leq \mu)$, so $y^{(\mu+1)}$ is continuous at $\xi_{\mu+1}$. But $y^{(\mu+2)}(t)$ is discontinuous at $\xi_{\mu+1}$, since $y^{(\mu+2)}(\xi_{\mu+1})$ includes $y^{(\mu+1)}(\xi_\mu)$, and $y^{(\mu+1)}$ is not continuous at $\xi_{\mu+1}$.

3.1.4 Delay VIDEs of neutral type with weakly singular kernels

Consider

$$\begin{cases} y'(t) = f(t, y(t), y(\theta(t))) + \displaystyle\int_0^t (t-s)^{-\alpha} G(s, y(s), y'(s), y'(\theta(s)))ds, & t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \tag{3.1.4.1}$$

where we assume $0 < \alpha < 1$ and (i), (ii), (iii) in (3.1.1.1).

We describe the primary discontinuities of problem (3.1.4.1) as the following theorem.

Theorem 3.1.4.1. *The primary discontinuities of problem (3.1.4.1) are generated inductively by the recursion (3.1.1.2), where $\xi_0 := 0$. More precisely, under the assumption* (i), (ii) *and* (iii) *in (3.1.1.1), $y \in C^{\mu,1-\alpha}(J^{[\mu]})$, but $y^{(\mu+1)}$ is not continuous at ξ_μ in general.*

Proof. (1) We consider the regularity of solution for (3.1.4.1) at the point $\xi_0 := 0$. It is possible to choose ϕ to satisfy $y(0) = \phi(0)$. However in general $\phi'(0-) \neq y'(0+)$.

Hence y is continuous at ξ_μ ($\mu \geq 0$), but y' is not continuous at ξ_0.

(2) Consider the regularity at ξ_1. Obviously, y' is continuous at ξ_1, hence y' is continuous at ξ_μ ($\mu \geq 1$). We can prove

$$H_1(t) := \int_0^t (t-s)^{-\alpha}\overline{G}(s)ds \in C^{1-\alpha}(J^{[1]}),$$

by following the lines proving (3.1.1.4).

(3) Consider the regularity at ξ_2. We write the equation (3.1.4.1) as

$$\begin{cases} y'(t) = f(t, y(t)) + \displaystyle\int_0^{\xi_1} (t-s)^{-\alpha}\overline{G}(s)ds + H_2(t), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases}$$

where $H_2(t) := \displaystyle\int_{\xi_1}^t (t-s)^{-\alpha}\overline{G}(s)ds$. By using the formula (3.1.1.3), we obtain

$$\begin{aligned} H_2(t) &:= \frac{1}{1-\alpha}\tilde{G}(\xi_1)(t-\xi_1)^{(1-\alpha)} + \frac{1}{1-\alpha}\int_{\xi_1}^t \overline{G}'(s)(t-s)^{1-\alpha}ds \\ H_2'(t) &= \overline{G}(\xi_1)(t-\xi_1)^{-\alpha} + \int_{\xi_1}^t \overline{G}'(s)(t-s)^{-\alpha}ds. \end{aligned}$$

Hence y'' is continuous at ξ_2. Thus y'' is continuous at ξ_μ ($\mu \geq 2$). Furthermore we can prove that $y'' \in C^{1-\alpha}(J^{[\mu]})$.

We consider the regularity at $\xi_{\mu+1}$. We rewrite the equation (3.1.4.1) as

$$\begin{cases} y'(t) = f(t, y(t)) + \displaystyle\int_0^{\xi_\mu} (\xi_\mu - s)^{-\alpha}\overline{G}(s)ds + H_{\mu+1}(t), \quad t \in I := [0, T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases}$$

where $H_{\mu+1}(t) := \int_{\xi_\mu}^t (t-s)^{-\alpha}\overline{G}(s)ds$.

$$\begin{aligned} H_{\mu+1}(t) &= \frac{1}{1-\alpha}\overline{G}(\xi_\mu)(t-\xi_\mu)^{1-\alpha} + \frac{1}{(1-\alpha)(2-\alpha)}\overline{G}'(\xi_\mu)(t-\xi_\mu)^{2-\alpha} \\ &+ \cdots + \frac{1}{(1-\alpha)_{\mu+1}}\overline{G}^{(\mu)}(\xi_\mu)(t-\xi_\mu)^{\mu+1-\alpha} \\ &+ \frac{1}{(1-\alpha)_{\mu+1}}\int_{\xi_\mu}^t \overline{G}^{(\mu+1)}(s)(t-s)^{\mu+1-\alpha}ds. \end{aligned} \tag{3.1.4.2}$$

Thus we can calculate the derivatives of order up to $\mu+1$ of $H_{\mu+1}(t)$ as:

$$\begin{aligned} H'_{\mu+1}(t) &= \overline{G}(\xi_\mu)(t-\xi_\mu)^{-\alpha} + \frac{1}{1-\alpha}\overline{G}'(\xi_\mu)(t-\xi_\mu)^{1-\alpha} \\ &+ \cdots + \frac{1}{(1-\alpha)_\mu}\overline{G}^{(\mu)}(\xi_\mu)(t-\xi_\mu)^{\mu-\alpha} \\ &+ \frac{1}{(1-\alpha)_\mu}\int_{\xi_\mu}^t \overline{G}^{(\mu+1)}(s)(t-s)^{\mu-\alpha}ds. \qquad (3.1.4.3) \\ &\vdots \\ H^{(\mu+1)}_{\mu+1}(t) &= (\alpha)_\mu\overline{G}(\xi_\mu)(t-\xi_\mu)^{-\alpha-\mu} + \cdots + \overline{G}^{(\mu)}(\xi_\mu)(t-\xi_\mu)^{-\alpha} \\ &+ \frac{1}{(1-\alpha)_\mu}\int_{\xi_\mu}^t \overline{G}^{(\mu+1)}(s)(t-s)^{-\alpha}ds. \qquad (3.1.4.4) \end{aligned}$$

Hence $y^{(\mu+1)}(t)$ is continuous at $\xi_{\mu+1}$. Furthermore we know that

$$\int_{\xi_\mu}^t \overline{G}^{(\mu+1)}(s)(t-s)^{-\alpha}ds \in C^{1-\alpha}(J^{[\mu+1]}),$$

thus $y \in C^{\mu+1,1-\alpha}(J^{[\mu+1]})$.

3.2 The discontinuous Galerkin method for delay VIDEs

3.2.1 The discontinuous Galerkin method for functional VIDEs of Hale's type

In this section we analyze the discontinuous Galerkin method for

$$\begin{cases} \dfrac{d}{dt}\left(y(t) - \displaystyle\int_0^{\theta(t)} K(t-s)G(y(s))ds\right) = f(t,y(t),y(\theta(t))), \quad t \in I := [0,T], \\ y(t) = \phi(t), \quad t \in [\bar{a}, 0], \end{cases} \qquad (3.2.1.1)$$

where we assume $y \in \Omega \subset \mathbb{R}$ and

(a) K is sufficiently smooth.

(b) (i), (ii), (iii) in (3.1.1.1).

Without loss of the generality we set $T = \xi_{M+1}$ for some $M \geq 1$, and introduce $Z_M := \{\xi_\mu : \ \mu = 0, 1, \dots, M\}$.

Since, as we have already seen in Section 3.1.2, the solution of (3.2.1.1) suffers from a loss of regularity at the primary discontinuity points $\{\xi_\mu\}$, the meshes I_h underlying the DG space will have to include these points if the DG solution is to attain its optimal global (or local) order. Thus, we shall employ meshes of the form

$$I_h := \bigcup_{\mu=0}^{M} I_h^{[\mu]}, \tag{3.2.1.2}$$

with the local mesh given by

$$I_h^{[\mu]} := \{t_n^{[\mu]} : \ \xi_\mu = t_0^{[\mu]} < t_1^{[\mu]} < \dots < t_{N_\mu}^{[\mu]} = \xi_{\mu+1}\} \ (\xi_\mu \in Z_M).$$

Such a mesh is called a *constrained mesh* (with respect to θ) for I. We introduce the following notations

$$I^{[\mu]} := (\xi_\mu, \xi_{\mu+1}], \ I_n^{[\mu]} := (t_{n-1}^{[\mu]}, t_n^{[\mu]}], \ \bar{I}^{[\mu]} := [\xi_\mu, \xi_{\mu+1}], \ \bar{I}_n^{[\mu]} := [t_{n-1}^{[\mu]}, t_n^{[\mu]}],$$

and

$$|I^{[\mu]}| := \xi_{\mu+1} - \xi_\mu, \ h_n^{[\mu]} := t_n^{[\mu]} - t_{n-1}^{[\mu]}, \ h^{[\mu]} := \max_{(n)} h_n^{[\mu]}.$$

Consider now the local *graded meshes* of the form

$$\begin{aligned} t_n^{[\mu]} &:= \left(\frac{n}{N_\mu}\right)^{r_\mu} \cdot |I^{[\mu]}| \\ &= \left(\frac{n}{N_\mu}\right)^{r_\mu} \cdot (\xi_{\mu+1} - \theta(\xi_{\mu+1})) \\ &= \left(\frac{n}{N_\mu}\right)^{r_\mu} \cdot \tau(\xi_{\mu+1}), \quad 0 \leq n \leq N_\mu - 1 \ (N_\mu \geq 2), \end{aligned} \tag{3.2.1.3}$$

where the *grading exponent* $r_\mu \in \mathbb{R}$ will always be assumed to satisfy $r_\mu \geq 1$. We know that

$$h_n^{[\mu]} \leq h^{[\mu]} \leq r_\mu \cdot \tau(\xi_{\mu+1}) N_\mu^{-1}, \quad 0 \leq n \leq N_\mu - 1 \ (N_\mu \geq 2).$$

In this presentation, for simplicity, we take $r_\mu = 1$ ($\forall \xi_\mu \in Z_M$) (i.e., each $I_h^{[\mu]}$ is a uniform mesh) and $N_\mu := (N)^{\frac{1}{\mu+1}}$ for all $\xi_\mu \in Z_M$.

We recall the problem (3.2.1.1) and denote

$$y^{[\mu]} := y|_{\bar{I}^{[\mu]}}.$$

Let

$$z(t) := y(t) - \int_0^{\theta(t)} K(t-s)G(y(s))ds. \tag{3.2.1.4}$$

It is easily seen from (3.2.1.4) that $z(t)$ and $y(t)$ possess the same regularity. We have

$$\begin{aligned} z'(t) &= f(t, y(t), y(\theta(t))) \\ &= f\left(t, z(t) + \int_0^{\theta(t)} K(t-s)G(y(s))ds, y(\theta(t))\right). \end{aligned} \tag{3.2.1.5}$$

We write (3.2.1.4) and (3.2.1.5) locally on the interval $I^{[\mu]}$.

$$y^{[\mu]} = z^{[\mu]} + \sum_{i=0}^{\mu-2} \int_{I^{[i]}} K(t-s)G(y^{[i]}(s))ds + \int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)G(y^{[\mu-1]}(s))ds. \tag{3.2.1.6}$$

$$\begin{aligned} \left(z^{[\mu]}\right)' &= f\left(t, z^{[\mu]} + \sum_{i=0}^{\mu-2} \int_{I^{[i]}} K(t-s)G(y^{[i]}(s))ds \right. \\ &+ \left. \int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)G(y^{[\mu-1]}(s))ds, y^{[\mu-1]}(\theta(t))\right). \end{aligned} \tag{3.2.1.7}$$

Here we set $y^{[0]} := \phi(0)$.

We abbreviate

$$\begin{aligned} \bar{f}(t) &:= f\left(t, z(t) + \int_0^{\theta(t)} K(t-s)G(y(s))ds, y(\theta(t))\right), \\ \bar{f}^{[\mu]} &:= f\left(t, z^{[\mu]} + \sum_{i=0}^{\mu-2} \int_{I^{[i]}} K(t-s)G(y^{[i]}(s))ds \right. \\ &+ \left. \int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)G(y^{[\mu-1]}(s))ds, y^{[\mu-1]}(\theta(t))\right). \end{aligned}$$

Define:

$$\mathcal{V}^{[\Lambda]}(I_h) := \left\{\varphi \in L^2(I) : \ \varphi|_{I_n^{[\mu]}} \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]}), \ \mu \in Z_M\right\},$$

where $\Lambda := \{\mu + 1\}_{\mu=0}^{M}$.

Now we are ready to define the $DG(\Lambda)$ scheme to (3.2.1.1): Find $z_h \in \mathcal{V}^{[\Lambda]}(I_h)$, such that

$$\sum_{\mu=0}^{M}\sum_{n=1}^{N_\mu}\left(\int_{I_n^{[\mu]}}[(z_h^{[\mu]})' - \bar{f}_h^{[\mu]}(t)]X(t)dt + [z_h^{[\mu]}]_n X_n^+\right) + (z_h^{[0]})_0^+ X_0^+ = \phi(0)X_0^+ \quad (3.2.1.8)$$

for all $X \in \mathcal{V}^{[\Lambda]}(I_h)$. Here,

$$\begin{aligned}
\bar{f}_h(t) &:= f\left(t, z_h(t) + \int_0^{\theta(t)} K(t-s)G(y_h(s))ds, y_h(\theta(t))\right), \\
\bar{f}_h^{[\mu]}(t) &:= f\left(t, z_h^{[\mu]} + \sum_{i=0}^{\mu-2}\int_{I^{[i]}} K(t-s)G(y_h^{[i]}(s))ds\right. \\
&+ \left.\int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)G(y_h^{[\mu-1]}(s))ds, y_h^{[\mu-1]}(\theta(t))\right).
\end{aligned}$$

This represents a time-stepping method: Find

$$z_h^{[\mu]} \in \mathcal{V}^{(\mu)}(I^{[\mu]}) := \left\{\varphi \in L^2(I^{[\mu]}) : \ \varphi|_{I_n^{[\mu]}} \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})\right\},$$

such that

$$\begin{aligned}
&\sum_{n=1}^{N_\mu}\int_{I_n{}^{[\mu]}}[(z_h^{[\mu]})' - \bar{f}_h^{[\mu]}(t)]X(t)dt + \sum_{n=2}^{N_\mu}[z_h^{[\mu]}]_{n-1}X_{n-1}^+ + (z_h^{[\mu]})_0^+ X_0^+ \\
= \ &(z_h^{[\mu]})_0^- X_0^+, \quad \forall X \in \mathcal{V}^{(\mu)}(I^{[\mu]}), \qquad (3.2.1.9)
\end{aligned}$$

for $\mu = 0, 1, \ldots, M$. Here, we set $(z_h^{[\mu]})_0^- := (z_h^{[\mu-1]})_{N_\mu}^-$ (for $\mu = 1, 2, \ldots, M$) and $(\xi^{[\mu]})_0^- = \phi(0)$ (for $\mu = 0$).

Also, (3.2.1.9) can be interpreted as a local time stepping method: Find $z_h^{[\mu]} \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$, such that

$$\int_{I_n^{[\mu]}}[(z_h^{[\mu]})' - \bar{f}_h^{[\mu]}]X(t)dt + (z_h^{[\mu]})_{n-1}^+ X_{n-1}^+ = (z_h^{[\mu]})_{n-1}^- X_{n-1}^+ \qquad (3.2.1.10)$$

for all $X \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$. Again, we set $(z_h^{[\mu]})_0^- := (z_h^{[\mu-1]})_{N_\mu}^-$. Hence define

$$y_h^{[\mu]} := z_h^{[\mu]} + \sum_{i=0}^{\mu-2} \int_{I^{[i]}} K(t-s)G(y_h^{[i]}(s))ds + \int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)G(y_h^{[\mu-1]}(s))ds. \quad (3.2.1.11)$$

We set

$$e_y := y - y_h,\ e_z := z - z_h,\ e_y^{[\mu]} := e_y|_{\bar{I}^{[\mu]}},\ e_z^{[\mu]} := e_z|_{\bar{I}^{[\mu]}}.$$

Subtracting (3.2.1.11) from (3.2.1.6) we obtain

$$\begin{aligned} e_y^{[\mu]} &:= e_z^{[\mu]} + \sum_{i=0}^{\mu-2} \int_{I^{[i]}} K(t-s)[G(y^{[i]}(s)) - G(y_h^{[i]}(s))]ds \\ &+ \int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)[G(y^{[\mu-1]}(s)) - G(y_h^{[\mu-1]}(s))]ds. \end{aligned}$$

Thus we have

$$||e_y^{[\mu]}||_{\bar{I}^{[\mu]}} \leq ||e_z^{[\mu]}||_{\bar{I}^{[\mu]}} + \sum_{i=0}^{\mu-2} \mathcal{K} L_G \left|I^{[i]}\right| ||e_y^{[i]}||_{\bar{I}^{[i]}} + \mathcal{K} L_G \left|I^{[\mu-1]}\right| ||e_y^{[\mu-1]}||_{\bar{I}^{[\mu-1]}},$$

where $\mathcal{K} := ||K||_{[\bar{a},T]}$, $L_f := ||f||_\Omega$, $L_G := ||G||_\Omega$. Hence, from Gronwall's lemma, we have

$$||e_y^{[\mu]}||_{\bar{I}^{[\mu]}} \leq ||e_z^{[\mu]}||_{\bar{I}^{[\mu]}} \exp(\mathcal{K} L_G \xi_\mu). \quad (3.2.1.12)$$

Now we estimate the error $||e_z^{[\mu]}||_{\bar{I}^{[\mu]}}$. Define the projection $\mathcal{I}z^{[\mu]} \in \mathcal{V}^{(\mu)}(I^{[\mu]})$, $\xi_\mu \in Z_M$, by

$$\begin{aligned} \left(\mathcal{I}z^{[\mu]}\right)_n^- &:= (z^{[\mu]})_n^-,\ 1 \leq n \leq N_\mu, & (3.2.1.13) \\ \int_{I_n^{[\mu]}} \left(\mathcal{I}z^{[\mu]}\right) X'(t)dt &:= \int_{I_n^{[\mu]}} z^{[\mu]} X'(t)dt & (3.2.1.14) \end{aligned}$$

for all $X \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$, $1 \leq n \leq N_\mu$. The approximation properties of $\mathcal{I}$ in (3.2.1.13) and (3.2.1.14) have been thoroughly investigated in [96]: On the generic subinterval $I_n^{[\mu]}$ there holds

$$||z^{[\mu]} - \mathcal{I}z^{[\mu]}||_{\bar{I}_n^{[\mu]}} \leq C(h_n^{[\mu]})^{\mu+1} ||(z^{[\mu]})^{(\mu+1)}||_{\bar{I}_n^{[\mu]}}, \quad (3.2.1.15)$$

where C is independent of $h_n^{[\mu]}$.

We split the error $e_z^{[\mu]} = z^{[\mu]} - z_h^{[\mu]} := \rho^{[\mu]} + \eta^{[\mu]}$ into $\rho^{[\mu]} := z^{[\mu]} - \mathcal{I}z^{[\mu]}$ and $\eta^{[\mu]} := \mathcal{I}z^{[\mu]} - z_h^{[\mu]}$. It is easily seen that $\eta^{[\mu]}$ satisfies

$$\int_{I_n^{[\mu]}} (\eta^{[\mu]})' X dt + (\eta^{[\mu]})_{n-1}^+ X_{n-1}^+ = \int_{I_n^{[\mu]}} [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)] X(t) dt + (\eta^{[\mu]})_{n-1}^- X_{n-1}^+, \tag{3.2.1.16}$$

for all $X \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$. Equivalently,

$$-\int_{I_n^{[\mu]}} (\eta^{[\mu]}) X' dt + (\eta^{[\mu]})_n^- X_n^- = \int_{I_n^{[\mu]}} [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)] X(t) dt + (\eta^{[\mu]})_{n-1}^- X_{n-1}^+. \tag{3.2.1.17}$$

for all $X \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$.

Lemma 3.2.1.1. *We have*

$$\begin{aligned}
[(\eta^{[\mu]})_n^-]^2 &\leq 2L_f \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt + (4L_f + L_f L_G \mathcal{K} \xi_\mu) \int_{I_n^{[\mu]}} [\eta^{[\mu]}]^2 dt \\
&+ L_f h_n^{[\mu]} ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + L_f L_G \mathcal{K} h_n^{[\mu]} \sum_{i=0}^{\mu-1} |I^{[i]}| ||e_y^{[i]}||^2_{\bar{I}^{[i]}} + [(\eta^{[\mu]})_{n-1}^-]^2.
\end{aligned}$$

Proof. We take $X = \eta^{[\mu]}$ in (3.2.1.16) and obtain

$$\begin{aligned}
&\frac{1}{2}[(\eta^{[\mu]})_n^-]^2 + \frac{1}{2}[(\eta^{[\mu]})_{n-1}^+]^2 \\
\leq\ & L_f \int_{I_n^{[\mu]}} \Bigg\{ \Bigg| e_z^{[\mu]} + \sum_{i=0}^{\mu-2} \int_{I^{[i]}} K(t-s)[G(y^{[i]}(s)) - G(y_h^{[\mu-1]}(s))] ds \\
+\ & \int_{\xi_{\mu-1}}^{\theta(t)} K(t-s)[G(y^{[\mu-1]}(s)) - G(y_h^{[\mu-1]}(s))] ds \Bigg| + |e_y^{[\mu-1]}(\theta(t))| \Bigg\} \\
& \cdot |\eta^{[\mu-1]}| dt + \frac{1}{2}[(\eta^{[\mu]})_{n-1}^-]^2 + \frac{1}{2}[(\eta^{[\mu]})_{n-1}^+]^2.
\end{aligned}$$

This yields

$$\begin{aligned}
[(\eta^{[\mu]})_n^-]^2 &\leq 2L_f \int_{I_n^{[\mu]}} |e_z^{[\mu]}||\eta^{[\mu]}| dt + 2L_f \int_{I_n^{[\mu]}} \left| e_y^{[\mu-1]}(\theta(t)) \right| \cdot |\eta^{[\mu]}| dt \\
&+ 2L_f L_G \mathcal{K} \int_{I_n^{[\mu]}} \left(\sum_{i=0}^{\mu-1} |I^{[i]}| \cdot ||e_y^{[i]}||_{\bar{I}^{[i]}} \right) |\eta^{[\mu]}| dt + [(\eta^{[\mu]})_{n-1}^-]^2.
\end{aligned}$$

Since $|e_z^{[\mu]}||\eta^{[\mu]}| \le |\rho^{[\mu]}|^2 + \frac{3}{2}|\eta^{[\mu]}|^2$, the lemma is proved.

Lemma 3.2.1.2. *We have*

$$
\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{n-1}^{[\mu]}) dt \\
\le\ & 6h_n^{[\mu]} L_f^2 \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt + 6h_n^{[\mu]} L_f^2 \int_{I_n^{[\mu]}} [\eta^{[\mu]}]^2 dt + 3(h_n^{[\mu]})^2 L_f^2 ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\
+\ & 3(h_n^{[\mu]})^2 L_f^2 L_G^2 \mathcal{K}^2 \mu \max_{i=0}^{\mu-1} |I^{[i]}|^2 \cdot \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}.
\end{aligned}
$$

Proof. To verify the lemma, we select $X = (\eta^{[\mu]})'(t - t_{n-1}^{[\mu]})$ in (3.2.1.17) and obtain

$$
\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{n-1}^{[\mu]}) dt \\
=\ & \int_{I_n^{[\mu]}} [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)] \left(\eta^{[\mu]}\right)' \cdot (t - t_{n-1}^{[\mu]}) dt \\
\le\ & \left(\int_{I_n^{[\mu]}} (t - t_{n-1}^{[\mu]}) [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)]^2 dt \right)^{1/2} \cdot \left(\int_{I_n^{[\mu]}} (t - t_{n-1}^{[\mu]}) \cdot [(\eta^{[\mu]})']^2 dt \right)^{1/2}.
\end{aligned}
$$

Hence we have

$$
\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{n-1}^{[\mu]}) dt \\
\le\ & h_n^{[\mu]} \int_{I_n^{[\mu]}} [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)]^2 dt \\
\le\ & h_n^{[\mu]} \int_{I_n^{[\mu]}} \left[L_f |e_z^{[\mu]}| + L_f |e_y^{[\mu-1]}(\theta(t))| + L_f L_G \mathcal{K} \sum_{i=0}^{\mu-1} |I^{[i]}| \cdot ||e_y^{[i]}||_{\bar{I}^{[i]}} \right]^2 dt \\
\le\ & 3h_n^{[\mu]} L_f^2 \int_{I_n^{[\mu]}} [e_z^{[\mu]}]^2 dt + 3(h_n^{[\mu]})^2 L_f^2 ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\
+\ & 3(h_n^{[\mu]})^2 L_f^2 L_G^2 \mathcal{K}^2 \mu \max_{i=0}^{\mu-1} |I^{[i]}|^2 \cdot \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}} \\
\le\ & 6h_n^{[\mu]} L_f^2 \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt + 6h_n^{[\mu]} L_f^2 \int_{I_n^{[\mu]}} [\eta^{[\mu]}]^2 dt + 3(h_n^{[\mu]})^2 L_f^2 ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\
+\ & 3(h_n^{[\mu]})^2 L_f^2 L_G^2 \mathcal{K}^2 \mu \max_{i=0}^{\mu-1} |I^{[i]}|^2 \cdot \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}},
\end{aligned}
\tag{3.2.1.18}
$$

thus completing the proof.

Lemma 3.2.1.3. *We have*

$$\begin{aligned}\left(\int_{I_n^{[\mu]}} \eta^{[\mu]}(t)\right)^2 &\leq 2(h_n^{[\mu]})^2[(\eta^{[\mu]})_n^-]^2 + 4(h_n^{[\mu]})^3 L_f^2 \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt \\ &+ 4(h_n^{[\mu]})^3 L_f^2 \int_{I_n^{[\mu]}} [\eta^{[\mu]}]^2 dt + \frac{4}{3}(h_n^{[\mu]})^6 L_f^2 ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\ &+ \frac{4}{3}(h_n^{[\mu]})^6 L_f^2 L_G^2 \mathcal{K}^2 \mu \max_{i=0}^{\mu-1} |I^{[i]}|^2 \cdot \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}.\end{aligned}$$

Proof. We choose $X = t_{n-1}^{[\mu]} - t$ in (3.2.1.17) to obtain

$$\int_{I_n^{[\mu]}} \eta^{[\mu]}(t)dt - h_n^{[\mu]}(\eta^{[\mu]})_n^- = \int_{I_n^{[\mu]}} [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)](t_{n-1}^{[\mu]} - t)dt.$$

Hence, by the Cauchy-Schwarz inequality, we have

$$\left(\int_{I_n^{[\mu]}} \eta^{[\mu]}(t)dt\right)^2 \leq 2(h_n^{[\mu]})^2[(\eta^{[\mu]})_n^-]^2 + 2\int_{I_n^{[\mu]}} [\bar{f}^{[\mu]}(t) - \bar{f}_h^{[\mu]}(t)]^2 dt \cdot \int_{I_n^{[\mu]}} (t_{n-1}^{[\mu]} - t)^2 dt.$$

Recalling the second inequality of (3.2.1.18) we finish the proof.

To derive the error estimates we need also the following two lemmas from [96].

Lemma 3.2.1.4. *There holds*

$$\int_{I_n^{[\mu]}} [\varphi(t)]^2 dt \leq \frac{1}{h_n^{[\mu]}} \left(\int_{I_n^{[\mu]}} \varphi(t)dt\right)^2 + \frac{1}{2}\int_{I_n^{[\mu]}} (t_n^{[\mu]} - t)(t - t_{n-1}^{[\mu]})[\varphi'(t)]^2 dt \quad (3.2.1.19)$$

for all $\varphi(t) \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$.

Lemma 3.2.1.5. *There holds*

$$||\varphi||^2_{\bar{I}_n^{[\mu]}} \leq C\log(\mu+2)\int_{I_n^{[\mu]}} [\varphi'(t)]^2 (t - t_{n-1}^{[\mu]})dt + C[\varphi_n^-]^2$$

for all $\varphi \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$, $\mu \geq 0$. *The constant* C *is independent of* $I_n^{[\mu]}$ *and* μ.

Now we are ready to prove the error estimates.

Corollary 3.2.1.1. *Let $\mathcal{K} := ||K||_{[\bar{a},T]}$, $L_f := ||f||_\Omega$ and $L_G := ||G||_\Omega$. Then, for $h_n^{[\mu]} L_f$ small enough, we have*

$$||e_z^{[\mu]}||_{\bar{I}^{[\mu]}} \leq C(L_f, L_G, \mathcal{K}, T, \mu)||z^{[\mu]} - \mathcal{I}z^{[\mu]}||_{\bar{I}^{[\mu]}}. \tag{3.2.1.20}$$

Proof. Combine Lemma 3.2.1.3 and Lemma 3.2.1.4 into

$$\begin{aligned}
&\left(\int_{I_n^{[\mu]}} \eta^{[\mu]}(t)dt\right)^2 \\
\leq\ & c(h_n^{[\mu]})^2[(\eta^{[\mu]})_n^-]^2 + c[h_n^{[\mu]}]^3 L_f^2 \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt \\
+\ & c[h_n^{[\mu]}]^2 L_f^2 \left(\int_{I_n^{[\mu]}} \eta^{[\mu]}(t)dt\right)^2 + c[h_n^{[\mu]}]^4 L_f^2 \int_{I_n^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{n-1}^{[\mu]})dt \\
+\ & c[h_n^{[\mu]}]^6 L_f^2 ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + c[h_n^{[\mu]}]^6 L_f^2 L_G^2 \mathcal{K}^2 \max_{i=0}^{\mu-1} |I^{[i]}|^2 \cdot \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}.
\end{aligned}$$

where c is a generic constant independent of any parameter. Hence for $h_n^{[\mu]} L_f$ small enough, we have

$$\begin{aligned}
\left(\int_{I_n^{[\mu]}} \eta^{[\mu]}(t)dt\right)^2 \leq\ & c(h_n^{[\mu]})^2[(\eta^{[\mu]})_n^-]^2 + c[h_n^{[\mu]}]^3 L_f^2 \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt \\
+\ & c[h_n^{[\mu]}]^4 L_f^2 \int_{I_n^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{n-1}^{[\mu]})dt + c[h_n^{[\mu]}]^6 L_f^2 ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\
+\ & c[h_n^{[\mu]}]^6 L_f^2 L_G^2 \mathcal{K}^2 \max_{i=0}^{\mu-1} |I^{[i]}|^2 \cdot \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}.
\end{aligned} \tag{3.2.1.21}$$

We combine Lemma 3.2.1.1, Lemma 3.2.1.2 and Lemma 3.2.1.4 to obtain

$$\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]}(t))']^2 (t - t_{n-1}^{[\mu]}) dt + [(\eta^{[\mu]})_n^-]^2 \\
\leq\ & cL_f \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt + cL_f \int_{I_n^{[\mu]}} [\eta^{[\mu]}]^2 dt + [(\eta^{[\mu]})_{n-1}^-]^2 \\
+\ & cL_f h_n^{[\mu]} ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + cL_f L_G \mathcal{K} T h_n^{[\mu]} \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}} \\
\leq\ & cL_f \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt + \frac{cL_f}{h_n^{[\mu]}} \left(\int_{I_n^{[\mu]}} \eta^{[\mu]} dt \right)^2 + cL_f h_n^{[\mu]} \int_{I_n^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{n-1}^{[\mu]}) dt \\
+\ & [(\eta^{[\mu]})_{n-1}^-]^2 + cL_f h_n^{[\mu]} ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + cL_f L_G \mathcal{K} T h_n^{[\mu]} \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}.
\end{aligned}$$

Using (3.2.1.21) we obtain

$$\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]}(t))']^2 (t - t_{n-1}^{[\mu]}) dt + [(\eta^{[\mu]})_n^-]^2 \\
\leq\ & cL_f \int_{I_n^{[\mu]}} [\rho^{[\mu]}]^2 dt + ch_n^{[\mu]} L_f \int_{I_n^{[\mu]}} [(\eta^{[\mu]}(t))']^2 (t - t_{n-1}^{[\mu]}) dt + cL_f h_n^{[\mu]} [(\eta^{[\mu]})_n^-]^2 \\
+\ & [(\eta^{[\mu]})_{n-1}^-]^2 + cL_f h_n^{[\mu]} ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + cL_f L_G \mathcal{K} T h_n^{[\mu]} \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}. \qquad (3.2.1.22)
\end{aligned}$$

Iterating (3.2.1.22) yields

$$\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]}(t))']^2 (t - t_{n-1}^{[\mu]}) dt + [(\eta^{[\mu]})_n^-]^2 \\
\leq\ & cL_f \sum_{i=1}^{n} h_i^{[\mu]} ||\rho^{[\mu]}||^2_{\bar{I}_i^{[\mu]}} + cL_f \sum_{i=1}^{n} \left(\int_{I_i^{[\mu]}} [(\eta^{[\mu]})']^2 (t - t_{i-1}^{[\mu]}) dt + [(\eta^{[\mu]})_i^-]^2 \right) \\
+\ & cL_f |I^{[\mu]}| \cdot ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + cL_f L_G \mathcal{K} T |I^{[\mu]}| \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}}.
\end{aligned}$$

For all sufficiently small $h_n^{[\mu]} L_f$, Gronwall's lemma can be applied and gives

$$\begin{aligned}
&\int_{I_n^{[\mu]}} [(\eta^{[\mu]}(t))']^2 (t - t_{n-1}^{[\mu]}) dt + [(\eta^{[\mu]})_n^-]^2 \\
\leq\ & \left(cL_f |I^{[\mu]}| \cdot ||\rho^{[\mu]}||^2_{\bar{I}^{[\mu]}} + cL_f |I^{[\mu]}| \cdot ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + cL_f L_G \mathcal{K} T |I^{[\mu]}| \sum_{i=0}^{\mu-1} ||e_y^{[i]}||^2_{\bar{I}^{[i]}} \right) \\
\cdot\ & \exp(cL_f |I^{[\mu]}|).
\end{aligned}$$

Observing Lemma 3.2.1.5 we then find

$$||\eta^{[\mu]}||^2_{\bar{I}_n^{[\mu]}} \leq \log(\max(\mu+1,2))cL_f|I^{[\mu]}| \left[||\rho^{[\mu]}||^2_{\bar{I}^{[\mu]}} + ||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} + L_G\mathcal{K}T\sum_{i=0}^{\mu-1}||e_y^{[i]}||^2_{\bar{I}^{[i]}} \right].$$

We obtain, by using the triangle inequality,

$$\begin{aligned} ||e_z^{[\mu]}||^2_{\bar{I}_n^{[\mu]}} &\leq C(L_f,T,\mu)||\rho^{[\mu]}||^2_{\bar{I}^{[\mu]}} + c\log(\max(\mu+1,2))L_f|I^{[\mu]}|\cdot||e_y^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\ &+ c\log(\max(\mu+1,2))L_fL_G\mathcal{K}T|I^{[\mu]}|\sum_{i=0}^{\mu-1}||e_y^{[i]}||^2_{\bar{I}^{[i]}}. \end{aligned}$$

From (3.2.1.12) we know that

$$\begin{aligned} ||e_z^{[\mu]}||^2_{\bar{I}_n^{[\mu]}} &\leq C(L_f,T,\mu)||\rho^{[\mu]}||^2_{\bar{I}^{[\mu]}} \\ &+ c\log(\max(\mu+1,2))L_f|I^{[\mu]}|\exp(2\mathcal{K}L_G\xi_{\mu-1})||e_z^{[\mu-1]}||^2_{\bar{I}^{[\mu-1]}} \\ &+ c\log(\max(\mu+1,2))L_fL_G\mathcal{K}T|I^{[\mu]}|\sum_{i=0}^{\mu-1}\exp(2\mathcal{K}L_G\xi_i)\cdot||e_z^{[i]}||^2_{\bar{I}^{[i]}}. \end{aligned}$$

We use Gronwall's lemma again and obtain

$$||e_z^{[\mu]}||_{\bar{I}^{[\mu]}} \leq C(L_f,L_G,\mathcal{K},T,\mu)||\rho^{[\mu]}||_{\bar{I}^{[\mu]}}.$$

Our main result is presented as the following theorem.

Theorem 3.2.1.1. *Under the assumptions in Corollary 3.2.1.1, we have*

$$||e_y||_I \leq C(L_f,L_G,\mathcal{K},T,\mu,M,y)\cdot\frac{1}{N}. \tag{3.2.1.23}$$

Proof. From Corollary 3.2.1.1 and (3.2.1.15) we have

$$||e_z^{[\mu]}||_{\bar{I}^{[\mu]}} \leq C(L_f,L_G,\mathcal{K},T,\mu,z)[h_n^{[\mu]}]^{\mu+1}. \tag{3.2.1.24}$$

Combining (3.2.1.24) and (3.2.1.12) we arrive at

$$||e_y^{[\mu]}||_{\bar{I}^{[\mu]}} \leq C(L_f,L_G,\mathcal{K},T,\mu,y)[h_n^{[\mu]}]^{\mu+1}. \tag{3.2.1.25}$$

We obtain

$$||e_y||_I \leq C(L_f,L_G,\mathcal{K},T,\mu,M,y)\cdot\frac{1}{N}.$$

3.2.2 The DG method for delay VIDEs of neutral type with weakly singular kernels

In this section we establish the convergence results of the DG method for equations with weakly singular kernels (3.1.2.3), (3.1.2.7), and (3.1.4.1).

In view of Theorem 3.1.2.4, the solution of (3.1.2.7) satisfies

$$y \in C^{1,1-\alpha}(J^{[\mu]}) \quad (\forall \mu \geq 1).$$

Hence, $DG(\Lambda)$ with $\Lambda := \{\mu + 1\}_{\mu=0}^{M}$ for equation (3.1.2.7) can only yield

$$||y - y_h||_{\bar{I}^{[\mu]}} = \mathcal{O}(N_\mu^{-(2-\alpha)}),$$

on uniform meshes. If we use the *graded meshes* with *grading exponent*

$$r_\mu := (\mu + 1)/(1 - \alpha)$$

(see (3.2.1.3)), then we may achieve the following theorem

Theorem 3.2.2.1. *The error estimate of $DG(\Lambda)$ method for equation (3.1.2.7) satisfies*

$$||y - y_h||_{\bar{I}^{[\mu]}} = \mathcal{O}(N_\mu^{-(\mu+2)}). \tag{3.2.2.1}$$

Proof. We can analyze the $DG(\Lambda)$ for equation (3.1.2.7) by following that of Section 3.2.1 except the estimate (3.2.1.15). The paper [96] shows that on the generic subinterval $I_n^{[\mu]}$ there holds

$$||z^{[\mu]} - \mathcal{I}z^{[\mu]}||_{\bar{I}_n^{[\mu]}} \leq C(\mu + 2)||z^{[\mu]} - q||_{L^2(\bar{I}_n^{[\mu]})} + C||(z^{[\mu]})' - q'||_{L^2(\bar{I}_n^{[\mu]})}, \tag{3.2.2.2}$$

for any $q \in \mathcal{P}^{(\mu+1)}(I_n^{[\mu]})$. Here and throughout this section, positive constant C is independent of $h_n^{[\mu]}$. Babuška and Suri [5] stated that

$$||z^{[\mu]} - \bar{z}||_{L^2(\bar{I}^{[\mu]})} \leq CN_\mu^{-(\mu+2)}, \tag{3.2.2.3}$$

and

$$||(z^{[\mu]})' - \bar{z}'||_{L^2(\bar{I}^{[\mu]})} \leq CN_\mu^{-(\mu+2)}, \tag{3.2.2.4}$$

where $\bar{z}$ is the FEM approximation to $z^{[\mu]}$ on the *graded meshes* (3.2.1.3) with *grading exponent*

$$r_\mu := (\mu + 1)/(1 - \alpha).$$

The original proof of (3.2.2.4) can be found in Gui and Babuška [54] (compare also Rice [93]).

Combining (3.2.2.2), (3.2.2.3), and (3.2.2.4) leads to

$$||z^{[\mu]} - \mathcal{I}z^{[\mu]}||_{\bar{I}^{[\mu]}} \leq CN_\mu^{-(\mu+2)}. \tag{3.2.2.5}$$

This estimate essentially helps us to complete the proof of (3.2.2.1).

The reads may compare Brunner [24, 23] which established these results in collocation methods on graded meshes for weakly singular Volterra integral and integro-differential equations.

Similarly, if we choose the *grading exponent* as

$$r_\mu = (\mu + 1)/(1 - \alpha),$$

then we may obtain the following theorem

Theorem 3.2.2.2. *We have the error estimate of DG*(Λ) *for equation (3.1.2.5)*

$$||y - y_h||_{\bar{I}^{[\mu]}} = \mathcal{O}(N_\mu^{-(2\mu+1)}).$$

Proof. The proof is similar to that of Theorem 3.2.2.1.

Remark 3.2.2.1. *As for DG*(Λ) *for equation (3.1.4.1), we can establish the convergence results by combining the techniques in Section 3.2.1 and [80].*

Chapter 4

Cascading multilevel discretization method for parabolic problems

4.1 Introduction

The two-grid method was first proposed by Xu [121, 120, 119] and later further studied by many others such as [4], [11], [38], [39], [81], [88], [116], [123], and [124]. So far the cascading multilevel discretization method has been investigated by [47], [71], [72], and [86].

The scheme of Marion and Xu [88] for the semi-linear parabolic equation (4.2.1.1) is based on two different finite element spaces one defined on a coarse grid with grid size H, and the other one on a fine grid with grid size $h \ll H$, respectively. Nonlinear and time-dependence are both treated in the coarse space, and only a fixed stationary equation needs to be solved on the fine space at each time. Howe.er a question arises: "how to solve the two equations efficiently?". We realize that when the coarse grid size is small, solving the nonlinear time-dependent equation on the coarse grid is not trivial. In [86] and Section 4.2 of this thesis, we construct cascading multilevel algorithms (Algorithm A and Algorithm B) based on the scheme proposed by Marion and Xu [88]. In Algorithm A, only fixed stationary linear equations need to be solved at each step and the result shows that the convergence rate is $O(h_J)$ in the energy norm $|\cdot|_1$, but it depends on the number of the grids. Algorithm B requires to solve both stationary linear equations and linear parabolic equations at each step, the total dimension of which equals to that of the corresponding level of the P1 conforming finite element space. The convergence rate of Algorithm B is also $O(h_J)$ in the energy norm $|\cdot|_1$ and is independent of the number of the grids. We also present Algorithm C, Algorithm D and Algorithm E for the parabolic equation with variable delays, parabolic equation with memory term and parabolic Fredholm equation.

4.2 Cascading multilevel discretization algorithms

In this section, we construct the cascading multilevel discretization algorithms and derive their convergence theorems.

4.2.1 Algorithm A

We consider the semilinear equation

$$u_t - \Delta u + f(u) = 0, \quad \text{in } \Omega \times \mathbb{R}^+, \tag{4.2.1.1}$$

with initial condition

$$u(x,0) = \hat{u}(x) \quad \text{in } \Omega,$$

and boundary condition

$$u = 0, \quad \text{on } \partial\Omega,$$

where $\Omega \subset \mathbb{R}^d$ (with $d \leq 3$) is a bounded convex polygonal domain. The nonlinear term f from $\mathbb{R}$ into $\mathbb{R}$ is assumed to be of class C^4 and its derivatives of order up to four are bounded on $\mathbb{R}$. We start our analysis from the weak form of (4.2.1.1):

$$(u_t, v) + ((u, v)) + (f(u), v) = 0, \ \forall v \in H_0^1(\Omega). \tag{4.2.1.2}$$

where $((u, v)) := (\nabla u, \nabla v)$, and $(\cdot, \cdot)$ is the L_2 inner product. We denote $|u|_1 := ((u, u))^{1/2}$, and $||u||_0 := (u, u)^{1/2}$. Let $\mathcal{T}^{h_j}$, $j = 0, 1, \ldots, J$, be the nested quasi-uniform triangulations of Ω, and let V_j be the corresponding P1 conforming finite element spaces. Thus

$$V_0 \subset V_1 \subset V_2 \cdots \subset V_J \subset H_0^1(\Omega).$$

We assume $h_j = h_{j-1}/2$ $(j = 1, \ldots, J)$, without loss of generality. The corresponding P1 conforming finite element approximation for (4.2.2.1) is: Find $u_j \in V_j$ such that

$$(u_{j,t}, v) + ((u_j, v)) + (f(u_j), v) = 0, \ \forall v \in V_j, \tag{4.2.1.3}$$

with $u_j(0) = Q_j \hat{u}$, where the operator Q_j is defined in (4.2.1.8) below. We present some results for the P1 conforming finite element approximation to (4.2.1.1). First,

$$|v|_1 \leq ch_j^{-1}||v||_0, \ ||v||_{L^\infty(\Omega)} \leq ch_j^{-d/p}||v||_{L^p(\Omega)}, \text{ for } 1 \leq p \leq \infty, \ \forall v \in V_j, \tag{4.2.1.4}$$

which are the well-known inverse inequalities (cf. Ciarlet [35], Brenner and Scott [22], Xu [122]). We have the following estimates:

$$|(u - u_j)(t)|_1 \leq Ch_j|\log h_j|^\sigma, \tag{4.2.1.5}$$

$$||(u - u_j)^{(i)}(t)||_0 \leq Ch_j^2|\log h_j|^\sigma, \ 0 \leq i \leq 2, \tag{4.2.1.6}$$

where σ is some nonnegative constant. Here and throughout this chapter C denotes the generic constant, which is independent of h_j and j, but may depend on t. The proofs of (4.2.1.5) and (4.2.1.6) with $i = 0$ can be found in Johnson *et al.*

[78] and in Crouzeix, Thomée, and Wahlbin [36]. Their techniques together with those in Thomée [113] can be extended in a straightforward manner to higher-order derivatives.

Now we consider the splitting of the space V_j:

$$V_j = V_{j-1} \ominus V_j^{j-1}, \text{ with } V_j^{j-1} := (I - Q_{j-1})V_j, \tag{4.2.1.7}$$

where $Q_{j-1} : L^2(\Omega) \longrightarrow V_{j-1}$ is the L^2 orthogonal projection into V_{j-1}, as defined by

$$(Q_{j-1}v, v_j) := (v, v_j), \ \forall v_j \in V_{j-1}. \tag{4.2.1.8}$$

Note also that V_{j-1} and V_j^{j-1} are orthogonal with respect to the scalar product $(\cdot, \cdot)$. We introduce the operator $R_{j-1}^j : V_j \longrightarrow V_j^{j-1}$ by setting

$$((R_{j-1}^j v_j, \chi)) := ((v_j, \chi)), \ \forall \chi \in V_j^{j-1}. \tag{4.2.1.9}$$

The cascading multilevel Algorithm A and Algorithm B associated with $(V_j, V_{j-1}, V_j^{j-1})$ consist of looking for an approximate solution

$$u^j := v^{j-1} + w^j. \text{ with } u^j \in V_j, \ v^{j-1} \in V_{j-1}, \ w^j \in V_j^{j-1} \tag{4.2.1.10}$$

for (4.2.2.1).

Algorithm A: Solve directly with respect to v^0

$$\begin{aligned} (v_t^0, \phi) + ((v^0, \phi)) + (f(v^0), \phi) &= 0, \quad \forall \phi \in V_0, \\ v^0(0) &= Q_0 \hat{u}. \end{aligned} \tag{4.2.1.11}$$

For $j = 1, \ldots, J$, solve the linear system of equations for w^j:

$$((v^{j-1} + w^j, \chi)) + (f(v^{j-1}), \chi) = 0, \quad \forall \chi \in V_j^{j-1}, \tag{4.2.1.12}$$

$$u^j = v^{j-1} + w^j, \tag{4.2.1.13}$$

$$v^j = u^j. \tag{4.2.1.14}$$

Note that equation (4.2.1.12) defines w^j uniquely in terms of v^{j-1}. We denote the corresponding mapping by

$$w^j = \Phi(v^{j-1}). \tag{4.2.1.15}$$

In order to establish the convergence of Algorithm A, we shall prove or cite a number of lemmas and assertions.

The following lemmas (Lemma 4.2.1.1 and Lemma 4.2.1.2) and assertions ((4.2.1.18)–(4.2.1.25)) and their proofs can be found in Marion and Xu [88].

Lemma 4.2.1.1. *Let* $[v, \phi]_1 := (((I - R_j^{j-1})v, (I - R_j^{j-1})\phi))$. *Then* v_{j-1} *satisfies*

$$(v_{j-1,t}, \phi) + [v_{j-1}, \phi]_1 = (u_{j,t}, R_j^{j-1}\phi) + (f(u_j), R_j^{j-1}\phi - \phi), \quad \forall \phi \in V_{j-1}. \tag{4.2.1.16}$$

Lemma 4.2.1.2. *Let* $||\phi||_1 := |(I - R_j^{j-1})\phi|_1$. *Then there exist two constants* C_1 *and* C_2, *independent of the grid size, such that,*

$$C_1|\phi|_1 \leq ||\phi||_1 \leq C_2|\phi|_1, \quad \forall \phi \in V_{j-1}. \tag{4.2.1.17}$$

For $t > 0$, we have

$$||w_j||_0 + ||w_{j,t}||_0 \leq Ch_{j-1}^2, \tag{4.2.1.18}$$

$$|w_j - \Phi(v_{j-1})|_1 \leq Ch_{j-1}^3, \tag{4.2.1.19}$$

$$||w_j - \Phi(v_{j-1})||_0 \leq Ch_{j-1}^4, \tag{4.2.1.20}$$

where Φ is as in (4.2.1.15). The following two assertions are essential in proving the convergence of Algorithm A and also Algorithm B for (4.2.1.1).

$$|w_j - w^j|_1 \leq Ch_{j-1}^3 + (1 + C_3h_j)|e_j|_1, \tag{4.2.1.21}$$

$$||w_j - w^j||_0 \leq Ch_{j-1}^4 + (1 + C_4h_j)|e_j|_1. \tag{4.2.1.22}$$

$e_j := v_{j-1} - v^{j-1}$. The following estimates are also helpful in the proof of the results in the thesis:

$$||(f(u_j) - f(v^{j-1}), R_j^{j-1}\phi)||_0 \le Ch_{j-1}||e_j||_0|\phi|_1 + Ch_{j-1}^3|\phi|_1, \tag{4.2.1.23}$$

$$||(f(u^j) - f(u_j), \phi)||_0 \le C\{||e_j||_0 + h_{j-1}|e_j|_1 + h_{j-1}^4\}||\phi||_0, \tag{4.2.1.24}$$

$$|(u_{j,t}, R_j^{j-1}\phi)|_1 \le Ch_{j-1}^3|\phi|_1. \tag{4.2.1.25}$$

Lemma 4.2.1.3. *We assume that*

$$\tilde{u}^j := \tilde{v}^{j-1} + \tilde{w}^j, \text{ with } \tilde{u}^j \in V_j,\ \tilde{v}^{j-1} \in V_{j-1},\ \tilde{w}^j \in V_j^{j-1}, \tag{4.2.1.26}$$

and

$$\begin{aligned} (\tilde{v}_t^{j-1}, \phi) + ((\tilde{v}^{j-1}, \phi)) + (f(\tilde{v}^{j-1}), \phi) &= 0, \quad \forall \phi \in V_{j-1}, \\ \tilde{v}^{j-1}(0) &= Q_{j-1}\hat{u}. \end{aligned} \tag{4.2.1.27}$$

$$((\tilde{v}^{j-1} + \tilde{w}^j, \chi)) + (f(\tilde{v}^{j-1}), \chi) = 0, \quad \forall \chi \in V_j^{j-1}. \tag{4.2.1.28}$$

In particular, let

$$\tilde{u}^0 = u^0,\ \tilde{v}^0 = v^0,\ \tilde{w}^1 = w^1, \tag{4.2.1.29}$$

and the initial value $\hat{u} \in L^2(\Omega)$ be given. Assume also that

$$h_j^2|\log h_j|^\sigma \le h_{j-1}^2, \tag{4.2.1.30}$$

where σ is as in (4.2.1.5) and (4.2.1.6). Then

$$|u_j - \tilde{u}^j|_1 \le C(K_1)^j h_j, \tag{4.2.1.31}$$

where $K_1 = 2(1 + C_3h_{j-1})$ and C_3 as in (4.2.1.21).

Proof. Firstly, we note that (4.2.1.30) is necessary for the proof of (4.2.1.18) (see Marion and Xu [88]). Let

$$\begin{aligned}
|u_j - \tilde{u}^j|_1 &= |u_j - \tilde{v}^{j-1} - \tilde{w}^j|_1 = |u_j - u_{j-1} + w_j - \tilde{w}^j - w_j|_1 \\
&\leq |w_j - \tilde{w}^j|_1 + |w_j|_1 + Ch_{j-1}. \qquad (4.2.1.32)
\end{aligned}$$

We estimate the term

$$\begin{aligned}
|w_j - \tilde{w}^j|_1 &\leq Ch_{j-1}^3 + (1 + C_3 h_{j-1})|v_{j-1} - \tilde{v}^{j-1}|_1 \\
&= Ch_{j-1}^3 + (1 + C_3 h_{j-1})|u_{j-1} - \tilde{u}^{j-1} + v_{j-1} - u_{j-1}|_1 \\
&= Ch_{j-1}^3 + (1 + C_3 h_{j-1})|u_{j-1} - \tilde{u}^{j-1} + v_{j-1} - u_j + u_j - u_{j-1}|_1 \\
&\leq (1 + C_3 h_{j-1})|u_{j-1} - \tilde{u}^{j-1}|_1 + |w_j|_1 + Ch_{j-1} + Ch_{j-1}^3. \qquad (4.2.1.33)
\end{aligned}$$

In view of (4.2.1.32) and (4.2.1.33), we have

$$|u_j - \tilde{u}^j|_1 \leq (1 + C_3 h_{j-1})|u_{j-1} - \tilde{u}^{j-1}|_1 + C|w_j|_1 + Ch_{j-1} + Ch_{j-1}^3.$$

Noting (4.2.1.18), the previous inequality arrives at

$$|u_j - \tilde{u}^j|_1 \leq K_1|u_{j-1} - \tilde{u}^{j-1}|_1 + Ch_{j-1} \leq C(K_1)^j h_j, \qquad (4.2.1.34)$$

which is (4.2.1.31).

Theorem 4.2.1.1. *The error estimate of Algorithm A for (4.2.1.1) is given by*

$$|u - u^J|_1 \leq C(2K_2)^J h_J, \qquad (4.2.1.35)$$

where $K_2 = 1 + Ch_{j-1}^2$.

Proof. From (4.2.1.27), we know that $\tilde{v}^{j-1} = u_{j-1}$, thus (4.2.1.28) becomes,

$$((u_{j-1} + \tilde{w}^j, \chi)) + (f(u_{j-1}), \chi) = 0, \quad \forall \chi \in V_j^{j-1}. \qquad (4.2.1.36)$$

And (4.2.1.12) is

$$((u^{j-1} + w^j, \chi)) + (f(u^{j-1}), \chi) = 0, \quad \forall \chi \in V_j^{j-1}. \tag{4.2.1.37}$$

Combining (4.2.1.36) and (4.2.1.37), we obtain

$$((\tilde{w}^j - w^j, \chi)) = -((u_{j-1} - u^{j-1}, \chi)) - (f(u_{j-1}) - f(u^{j-1}), \chi). \tag{4.2.1.38}$$

substituting $\chi = \tilde{w}^j - w^j$ into (4.2.1.38), we get

$$\begin{aligned} |\tilde{w}^j - w^j|_1^2 &= -((u_{j-1} - u^{j-1}, \tilde{w}^j - w^j)) - (f(u_{j-1}) - f(u^{j-1}), \tilde{w}^j - w^j) \\ &\leq |u_{j-1} - u^{j-1}|_1 |\tilde{w}^j - w^j|_1 \\ &+ Ch_{j-1}^2 |u_{j-1} - u^{j-1}|_1 |\tilde{w}^j - w^j|_1. \end{aligned} \tag{4.2.1.39}$$

Hence,

$$|\tilde{w}^j - w^j|_1 \leq (1 + Ch_{j-1}^2)|u_{j-1} - u^{j-1}|_1. \tag{4.2.1.40}$$

Now we estimate

$$\begin{aligned} |u_j - u^j|_1 &= |u_j - \tilde{u}^j + \tilde{u}^j - u^j|_1 \\ &= |u_j - \tilde{u}^j + u_{j-1} - u^{j-1} + \tilde{w}^j - w^j|_1 \\ &\leq |u_j - \tilde{u}^j|_1 + |u_{j-1} - u^{j-1}|_1 + |\tilde{w}^j - w^j|_1 \\ &\leq C(K_1)^j h_j + |u_{j-1} - u^{j-1}|_1 + (1 + Ch_{j-1}^2)|u_{j-1} - u^{j-1}|_1 \\ &\leq C(K_1)^j h_j + K_2 |u_{j-1} - u^{j-1}|_1, \end{aligned} \tag{4.2.1.41}$$

where the second step follows from (4.2.1.13) and (4.2.1.26), and the fourth step is based on (4.2.1.31) and (4.2.1.40). Therefore, (4.2.1.41) yields

$$\begin{aligned} |u_J - u^J|_1 &\leq C(K_1)^J h_J + K_2 |u_{J-1} - u^{J-1}|_1 \\ &\leq C(K_1)^J h_J + C K_2 K_1{}^{J-1} h_{J-1} + K_2^2 \| u_{J-2} - u^{J-2} \| \\ &\vdots \\ &\leq C(K_1)^J h_J + C K_2 K_1^{J-1} h_{J-1} + \cdots + C K_2^J h_0 \\ &\leq (2K_2)^J h_J. \end{aligned} \tag{4.2.1.42}$$

Consequently, we achieve

$$|u - u^J|_1 = |u - u_J + u_J - u^J|_1 \leq |u - u_J|_1 + |u_J - u^J|_1 \leq C(2K_2)^J h_J,$$

which is (4.2.1.35).

Remark 4.2.1.1. *We know from Theorem 4.2.1.1 that when we fix the number of grids, the convergence rate of Algorithm A is* $O(h_J)$.

4.2.2 Algorithm B

In this section, we construct Algorithm B, which solves both time-dependent linear and stationary linear equations at each level of the P1 finite element space.

Algorithm B: Let

$$u^0 = Q_0 \hat{u}. \tag{4.2.2.1}$$

For $j = 2, \ldots, J$, solve

$$(v_t^{j-1}, \phi) + ((v^{j-1} + w^j, \phi)) + (f(u^{j-1}), \phi) = 0, \quad \forall \phi \in V_{j-1}, \tag{4.2.2.2}$$

$$((v^{j-1} + w^j, \chi)) + (f(v^{j-1}), \chi) = 0, \quad \forall \chi \in V_j^{j-1}, \tag{4.2.2.3}$$

$$u^j = v^{j-1} + w^j, \tag{4.2.2.4}$$

for v^{j-1} and w^j.

Theorem 4.2.2.1. *Under the assumptions of Theorem 2.1 we have*

$$|u - u_J|_1 \leq Ch_J, \tag{4.2.2.5}$$

where the constant C is independent of the grid size h_J and the grid J.

Proof. We use the corresponding notation of Section 4.2.2.1 and we claim that

$$(v_t^{j-1}, \phi) + [v^{j-1}, \phi]_1 = -(f(u^{j-1}), \phi) + (f(v^{j-1}), R_j^{j-1}\phi), \quad \forall \phi \in V_{j-1}. \tag{4.2.2.6}$$

The proof of (4.2.2.6) is the analog of that of (4.2.1.16). Because of (4.2.2.6) and (4.2.1.16), $d_j := v_{j-1} - v^{j-1}$ satisfies

$$(d_{j,t}, \phi) + [d_j, \phi]_1 = (u_{j,t}, R_j^{j-1}) + (f(u_j) - f(v^{j-1}), R_j^{j-1}\phi) + (f(u^{j-1}) - f(u_j), \phi). \tag{4.2.2.7}$$

We estimate

$$\begin{aligned} ||(f(u^{j-1}) - f(u_j), \phi)||_0 &\leq C||u^{j-1} - u_j||_0||\phi||_0 \\ &\leq (||u^{j-1} - u_{j-1}||_0 + Ch_{j-1})||\phi||_0. \end{aligned} \tag{4.2.2.8}$$

Substituting $\phi = d_{j,t}$ into (4.2.2.7) and using (4.2.2.8), we see that

$$\begin{aligned} ||d_{j,t}||_0^2 + [d, d_{j,t}]_1 &\leq (u_{j,t}, R_j^{j-1}d_{j,t}) + (f(u_j) - f(v^{j-1}), R_j^{j-1}d_{j,t}) \\ &+ (||u^{j-1} - u_{j-1}||_0 + Ch_{j-1})||\phi||_0. \end{aligned} \tag{4.2.2.9}$$

The estimates (4.2.1.23), (4.2.1.25), (4.2.1.17) and Young's inequality leads to

$$\frac{d(|d_j|_1^2)}{dt} \leq Ch_{j-1}^2|d_j|_1^2 + Ch_{j-1}^2|u^{j-1} - u_{j-1}|_1^2 + Ch_{j-1}^2.$$

Hence,

$$|d_j|_1^2 \leq Ch_{j-1}^2 + Ch_{j-1}^2|u^{j-1} - u_{j-1}|_1^2 + Ch_{j-1}^4 \int_0^t |u^{j-1} - u_{j-1}|_1^2 ds.$$

Thus we have

$$
\begin{aligned}
|u_j - u^j|_1^2 &\leq 2|d_j|_1^2 + 2|w_j - w^j|_1^2 \leq C|d_j|_1^2 + Ch_{j-1}^6 \\
&\leq Ch_{j-1}^2 + Ch_{j-1}^2|u_{j-1} - u^j|_1^2 + Ch_{j-1}^4 \int_0^t |u_{j-1} - u^{j-1}|_1^2 ds.
\end{aligned}
$$

Therefore,

$$
|u_j - u^j|_1 \leq Ch_{j-1} + Ch_{j-1}|u_{j-1} - u^j|_1. \tag{4.2.2.10}
$$

Iterating (4.2.2.10) yields

$$
\begin{aligned}
|u_J - u^J|_1 &\leq Ch_{J-1} + Ch_{J-1}|u_{J-1} - u^{J-1}|_1 \\
&\leq Ch_{J-1} + Ch_{J-1}h_{J-2} + Ch_{J-1}h_{J-2}|u_{J-2} - u^{J-2}|_1 \\
&\vdots \\
&\leq Ch_{J-1} + Ch_{J-1}h_{J-2} + \cdots + Ch_{J-1}h_{J-2}\cdots h_1 \\
&\leq Ch_J[1 + h_1(2^{-(J-3)} + 2^{-(J-3)}2^{-(J-4)} + \cdots \\
&+ 2^{-(J-3)}2^{-(J-4)}\cdots 1)] \leq Ch_J(1 + Ch_1) \leq Ch_J.
\end{aligned}
$$

Remark 4.2.2.1. *In Algorithm A and Algorithm B, we need to solve the linear equations (4.2.1.12) and (4.2.2.3), respectively. For the further study of the linear systems in* V_j^{j-1}*, the author refers to Marion and Xu [88].*

4.3 Comparison and discussion

We summarize the presentation of Section 4.2 in Table 4.3, and add some discussions.

Table 4.1: Comparison of Algorithm A and Algorithm B

Algorithm	Convergence	Open question	Literature
A	$O(h_J)$ (but dependent on J)	CMGI	[88]
B	$O(h_J)$ (independent on J)	CMGI	

Algorithm A and Algorithm B are based on the L_2 decomposition techniques. Algorithm A needs only to solve a linear fixed stationary system of equations in V_j^{j-1} of each step. Algorithm B requires the solution of the linear parabolic equations in V_{j-1} and the linear fixed stationary equations in V_j^{j-1} at each step. If we incorporate the classical iteration methods (called smoothers) into the linear stationary system or the linear system arising from the linear parabolic equations with discontinuous Galerkin (DG) time-stepping methods or other methods, then it is not difficult to formulate the cascading multigrid iteration method (CMGI), which is ongoing work. For the idea of cascading multigrid methods, Algorithm A avoids the solving of the linear parabolic equations. Hence, it is much easier to implement cascading multigrid iteration methods, but it can be used only when we fix the number of grids. Although Algorithm A and Algorithm B have the same convergence rate $O(h_J)$, Algorithm A is dependent of the number of grids J, while Algorithm B is not. So Algorithm B is more accurate than Algorithm A with respect to convergence: it has a smaller error constant.

4.4 Extensions to other parabolic problems

In this section we extend the analysis of the cascading multilevel discretization algorithms to parabolic partial differential equations with variable delays and with nonlinear memory terms.

4.4.1 Parabolic equation with delay argument

We consider

$$u_t(x,t) - \triangle u(x,t) + f(u(x,t), u(x,\theta(t))) = 0, \quad \forall x \in \Omega, \ t \in I := [0,T], \qquad (4.4.1.1)$$

with initial condition

$$u(x,t) = \hat{u}(x,t), \quad \forall x \in \Omega, \ \forall t \leq 0,$$

and boundary condition

$$u(x,t)|_{\partial\Omega} = 0, \quad \forall t \in I,$$

where $\Omega \subset \mathbb{R}^d$ (with $d \leq 3$) is a bounded convex polygonal domain. Consider the case of

$$f(u,v) = f_1(u) + f_2(v).$$

Here f_1 and f_2 from $\mathbb{R}$ into $\mathbb{R}$ are assumed to be of class C^1 and their derivatives are bounded on $\mathbb{R}$. The delay function $\theta(t)$ will be subject to the following conditions (i)–(iii):

(i) $\theta(t) = t - \tau(t), \ \theta \in C^d(I)$ for some $d \geq 0$;

(ii) $\tau(t) \geq \tau_0 > 0$ for $t \in I$;

(iii) θ is strictly increasing on I.

We define the points $\{\xi_\mu\}, \ \mu = 0, 1, \ldots, M$, by

$$\theta(\xi_\mu) = \xi_{\mu-1}, \quad \mu \geq 1,$$

where $\xi_0 := 0$. $I^{[i]} := [\xi_{i-1}, \xi_i], \ i = 1, \ldots, M$. Furthermore we assume, without loss of generality, that $T = \sum_{i=1}^{M} |I^{[i]}|$. We note that we shall use the corresponding notations in Section 4.2.2.1. We shall analyze the P1 conforming finite element approximation to (4.4.1.1). The weak form of (4.4.1.1) is

$$(u_t, v) + ((u,v)) + (f(u(x,t), u(x,\theta(t))), v) = 0, \quad \forall v \in H_0^1(\Omega). \tag{4.4.1.2}$$

If we set $u^{[i]} := u(x,t),\ t \in I^{[i]}$, then (4.4.1.2) can be written as

$$(u_t^{[i]}, v) + ((u^{[i]}), v)) + (f_1(u^{[i]}(x,t)), v) + (f_2(u^{[i-1]}(x,\theta(t))), v) = 0,\ \forall v \in H_0^1(\Omega). \tag{4.4.1.3}$$

for all $t \in I^{[i]}$. Correspondingly, the P1 conforming finite element approximation to (4.4.1.1) is

$$(u_{j,t}, v) + ((u_j, v)) + (f(u_j(x,t), u_j(x,\theta(t))), v) = 0, \tag{4.4.1.4}$$

for all $v \in V_j$ with $u_j := R_j\hat{u}(x,t)$ $(t \le 0)$, and

$$(u_{j,t}^{[i]}, v) + ((u_j^{[i]}, v)) + (f_1(u_j^{[i]}(x,t)), v) + (f_2(u_j^{[i-1]}(x,\theta(t))), v) = 0, \tag{4.4.1.5}$$

for all $v \in V_j$ $(t \in I^{[i]})$, where $u_j^{[i]} := u_j(x,t)$ $(t \in I^{[i]})$. We are now ready to present the convergence results of the P1 conforming finite element method for (4.4.1.1).

Theorem 4.4.1.1. *Let u and u_j be the solution of (4.4.1.1) and (4.4.1.4), respectively. Then we have*

$$\begin{aligned} ||u - u_j||_0 &\le Ch_j^2 |\log h_j|^\sigma, \\ |u - u_j|_1 &\le Ch_j |\log h_j|^\sigma, \end{aligned}$$

where σ is as in (4.2.1.5) and (4.2.1.6).

Proof. For convenience, we define

$$e(t) := u(x,t) - u_j(x,t),\ e^{[i]} := u^{[i]}(x,t) - u_j^{[i]}(x,t),\ \tilde{e}_j^{[i]} := \tilde{u}_j^{[i]} - u_j^{[i]},$$

where $\tilde{u}_j^{[i]}$ is the solution of

$$(\tilde{u}_{j,t}^{[i]}, v) + ((\tilde{u}_j^{[i]}, v)) + (f_1(\tilde{u}_j^{[i]}(x,t), v) + (f_2(u^{[i-1]}(x,\theta(t))), v) = 0,\ \forall v \in V_j, \tag{4.4.1.6}$$

for $t \in I^{[i]}$. From (4.2.1.5) and (4.2.1.6), we know that

$$
\begin{aligned}
||u^{[i]} - \tilde{u}_j^{[i]}||_0 &\leq Ch_j^2|\log h_j|^\sigma, \quad \forall t \in I^{[i]}, && \text{(4.4 1.7)} \\
|u^{[i]} - \tilde{u}_j^{[i]}|_1 &\leq Ch_j|\log h_j|^\sigma, \quad \forall t \in I^{[i]}. && \text{(4.4.1.8)}
\end{aligned}
$$

Because of

$$
e^{[i]}(t) := u^{[i]}(x,t) - u_j^{[i]}(x,t) = (u^{[i]} - \tilde{u}_j^{[i]}) + (\tilde{u}_j^{[i]} - u_j^{[i]}), \tag{4.4.1.9}
$$

we need only to estimate the term $\tilde{e}_j^{[i]} := \tilde{u}_j^{[i]} - u_j^{[i]}$. Subtracting (4.4.1.5) from (4.4.1.6), we obtain

$$
\begin{aligned}
&(\tilde{e}_{j,t}^{[i]}, v) + ((\tilde{e}_j^{[i]}, v)) + (f_1(\tilde{u}_j^{[i]}(x,t)) - f_1(u_j^{[i]}(x,t)), v) \\
&+(f_2(u^{[i-1]}(x,\theta(t))) - f_2(u_j^{[i-1]}(x,\theta(t))), v) = 0, \quad \forall v \in V_j. \quad \text{(4.4.1.10)}
\end{aligned}
$$

Substituting $v := e_j^{[i]}$ into (4.4.1.10) leads to

$$
\frac{1}{2}\frac{d\left(||\tilde{e}_j^{[i]}||_0^2\right)}{dt} + |\tilde{e}_j^{[i]}|_1^2 \leq C||\tilde{e}_j^{[i]}||_0^2 + C||e^{[i-1]}(\theta(t))||_0||\tilde{e}_j^{[i]}||_0,
$$

and

$$
\frac{d\left(||\tilde{e}_j^{[i]}||_0^2\right)}{dt} \leq C||\tilde{e}_j^{[i]}(t)||_0^2 + ||e^{[i-1]}(\theta(t))||_0^2.
$$

Consequently, we have

$$
||\tilde{e}_j^{[i]}||_0^2 \leq C||\tilde{e}_j^{[i]}(t_{i-1})||_0^2 + \int_{t_{i-1}}^{t} \exp(C\cdot(t-s))||e^{[i-1]}(\theta(s))||_0^2 ds.
$$

Hence, the estimates

$$
\begin{aligned}
||\tilde{e}_j^{[i]}(t)||_0 &\leq C||e^{[i-1]}(t)||_0, && \text{(4.4.1.11)} \\
|\tilde{e}_j^{[i]}(t)|_1 &\leq C|e^{[i-1]}(t)|_1, && \text{(4.4.1.12)}
\end{aligned}
$$

hold. According to (4.4.1.9), (4.4.1.7) and (4.4.1.8), we get

$$
||e^{[i]}(t)||_0 \leq ||u^{[i]}(t) - \tilde{u}_j^{[i]}(t)||_0 + ||\tilde{e}_j^{[i]}||_0 \leq Ch_j^2|\log h_j|^\sigma + C||e^{[i-1]}(t)||_0, \quad \forall t \in I^{[i]}.
$$

Similarly, we find

$$|e^{[i]}(t)|_1 \le Ch_j |\log h_j|^\sigma + C|e^{[i-1]}(t)|_1, \quad \forall t \in I^{[i]},$$

and this leads to

$$\begin{aligned} ||e(t)||_0 &\le Ch_j^2 |\log h_j|^\sigma, \\ |e(t)|_1 &\le Ch_j |\log h_j|^\sigma. \end{aligned}$$

Algorithm C: Let

$$u^0 = R_0 \hat{u}(x,t), \quad \forall t \le 0. \tag{4.4.1.13}$$

For $j = 2, \ldots, J$, solve

$$(u_t^j, v) + ((u^j, v)) + (f(u^{j-1}(x,t), u^{j-1}(x, \theta(t))), v) = 0, \quad \forall v \in V_j, \tag{4.4.1.14}$$

for u^j.

Theorem 4.4.1.2. *Assume that*

$$|\log h_{j-1}|^\sigma \le C. \tag{4.4.1.15}$$

Let u and u^J denote the solution of (4.4.1.1) and Algorithm C, respectively. Then we have

$$|u - u^J|_1 \le Ch_J,$$

where C is independent of h_J and the number of the grids.

Proof. Let $\delta_j := u_j - u^j$. Then subtracting (4.4.1.14) from (4.4.1.4), we get

$$(\delta_{j,t}, v) + ((\delta_j, v)) + (f(u_j(x,t), u_j(x, \theta(t))) - f(u^{j-1}(x,t), u^{j-1}(x, \theta(t))), v) = 0. \tag{4.4.1.16}$$

Now we estimate

$$
\begin{aligned}
& ||(f(u_j(x,t),u_j(x,\theta(t))) - f(u^{j-1}(x,t),u^{j-1}(x,\theta(t))),v)||_0 \quad (4.4.1.17)\\
\leq\ & ||(f(u_j(x,t),u_j(x,\theta(t))) - f(u^{j-1}(x,t),u^{j-1}(x,\theta(t)))||_0||v||_0\\
\leq\ & [C||u_j(x,t)-u_{j-1}(x,t)+\delta_{j-1}||_0\\
+\ & C||u_j(x,\theta(t))-u_{j-1}(x,\theta(t))+\delta_{j-1}(x,\theta(t))||_0]\cdot||v||_0\\
\leq\ & [Ch_{j-1}^2|\log h_{j-1}|^{2\sigma}+||\delta_{j-1}(x,t)||_0+||\delta_{j-1}(x,\theta(t))||_0]\cdot||v||_0.
\end{aligned}
$$

Substituting $v=\delta_{j,t}$ into (4.4.1.16) and combining (4.4.1.17), we obtain the estimate

$$
\begin{aligned}
||\delta_{j,t}||_0^2+\frac{d(|\delta_j|_1^2)}{dt} &\leq [Ch_{j-1}^2|\log h_{j-1}|^{2\sigma}+||\delta_{j-1}(x,t)||_0+||\delta_{j-1}(x,\theta(t))||_0]\cdot||\delta_{j,t}||_0\\
&\leq Ch_{j-1}^4|\log h_{j-1}|^{4\sigma}+h_{j-1}^2|\delta_{j-1}(x,t)|_1^2+h_{j-1}^2|\delta_{j-1}(x,\theta(t))|_1^2.
\end{aligned}
$$

Hence, it holds that

$$
|\delta_j|_1^2\leq Ch_{j-1}^4|\log h_{j-1}|^{4\sigma}+h_{j-1}^2\int_0^t(|\delta_{j-1}(x,s)|_1^2+|\delta_{j-1}(x,\theta(s))|_1^2)ds. \quad (4.4.1.18)
$$

Iterating (4.4.1.18) leads to

$$
\begin{aligned}
|\delta_j|_1^2 &\leq Ch_{j-1}^4|\log h_{j-1}|^{4\sigma}+Ch_{j-1}^2h_{j-2}^4|\log h_{j-2}|^{4\sigma}+\cdots\\
&+ Ch_{j-1}^2h_{j-2}^2\cdots h_1^4|\log h_1|^{4\sigma}\leq Ch_{j-1}^4. \quad (4.4.1.19)
\end{aligned}
$$

In the last step of (4.4.1.19), we used the assumption (4.4.1.15). Recalling Theorem 4.4.1.1 and (4.4.1.19), we obtain

$$
|u-u^J|_1\leq|u-u_J|_1+|u_J-u^J|_1\leq ch_J|\log h_J|^\sigma+ch_J\leq ch_J,
$$

which is our desired result.

4.4.2 Parabolic equation with memory term

We consider

$$u_t - \triangle u + \int_0^t k(t-s)G(u(x,s))ds = 0, \quad \forall x \in \Omega, \ \forall t \in I := [0,T], \tag{4.4.2.1}$$

with initial condition

$$u(x,0) = \hat{u}(x), \quad \forall x \in \Omega,$$

and boundary condition

$$u(x,t) = 0, \quad \forall x \in \partial\Omega, \ \forall t \in I,$$

where $\Omega \subset \mathbb{R}^d$ (with $d \leq 3$) is a bounded convex polygonal domain. The nonlinear term G from $\mathbb{R}$ into $\mathbb{R}$ is assumed to be of class C^1, with bounded derivatives on $\mathbb{R}$. Furthermore we assume $k \in C^1(I)$. The weak form of (4.4.2.1) is

$$(u_t, v) + ((u,v)) + \int_0^t k(t-s)(G(u(x,s)),v)ds = 0, \quad \forall v \in H_0^1(\Omega),$$

where $(G(u(x,s)),v) := \int_\Omega G(u(x,s))vdx$. Correspondingly, the P1 conforming finite element approximation is given by

$$(u_{j,t}, v) + ((u_j, v)) + \int_0^t k(t-s)(G(u_j(s)),v)ds = 0, \quad \forall v \in V_j, \tag{4.4.2.2}$$

with $u_j(0) = R_j\hat{u}$. We know that

$$|u - u_j|_1 \leq Ch_j, \tag{4.4.2.3}$$

which can be found in [34].

Algorithm D: Let

$$u^0 = R_0\hat{u}.$$

For $j = 2, \ldots, J$, solve

$$(u_t^j, v) + ((u^j, v)) + \int_0^t k(t-s)(G(u^{j-1}) + G'(u^{j-1})(u^j - u^{j-1}), v)ds = 0, \quad \forall v \in V_j, \tag{4.4.2.4}$$

for u^j.

Theorem 4.4.2.1. *Assume that*

$$h_{j-1} \leq \frac{C}{(2\sqrt{2})^j}. \tag{4.4.2.5}$$

Then

$$|u - u^J|_1 \leq Ch_J,$$

where C is independent on h_J and the number of the grids.

Proof. Let $e_j := u_j - u^j$. Subtracting (4.4.2.4) from (4.4.2.2) yields

$$\begin{aligned}(e_{j,t}, v) + ((e_j, v)) + \int_0^t k(t-s)\big(G(u_{j-1}) - G(u^{j-1}) + G'(u_{j-1})(u_j - u_{j-1})\\ -G'(u^{j-1})(u^j - u^{j-1}) + O((u_j - u_{j-1})^2), v\big)\, ds = 0, \quad \forall v \in V_j.\end{aligned} \tag{4.4.2.6}$$

Bringing $v = e_{j,t}$ into (4.4.2.6), we arrive at

$$||e_{j,t}||_0^2 + \frac{1}{2}\frac{d(|e_j|_1^2)}{dt} \leq Ch_{j-1}^4 + C(\int_0^t |k(t-s)||e_{j-1}|ds)^2.$$

Hence, we achieve

$$\frac{d(|e_j|_1^2)}{dt} \leq Ch_{j-1}^4 + Ch_{j-1}^2(\int_0^t |k(t-s)||e_{j-1}|_1 ds)^2 \leq C(h_{j-1}^4 + 2\int_0^t |e_{j-1}|_1^2 ds), \tag{4.4.2.7}$$

by using $|ab| \leq 2(a^2 + b^2)$.

Integrating (4.4.2.7) yields

$$|e_j|_1^2 \leq Ch_{j-1}^4 + 2t\int_0^t |e_{j-1}|_1^2 ds \leq Ch_{j-1}^4 + C2h_{j-2}^4 + \cdots + C2^{j-2}h_1^4 \leq Ch_j^2 8^j h_{j-1}^2.$$

Therefore, by using assumption (4.4.2.5) we get

$$|e_j|_1 \leq Ch_j \quad (j = 1, \ldots, J). \tag{4.4.2.8}$$

Then (4.4.2.8) and (4.4.2.3) lead to

$$|u - u^J|_1 \leq |u - u_J|_1 + |u_J - u^J|_1 \leq ch_J,$$

which completes the proof.

4.5 Application to parabolic Fredholm equation

4.5.1 Finite element method for parabolic Fredholm equation

Consider the parabolic Fredholm equation

$$u_t - \Delta u = \int_\Omega f(u)dx, \quad \forall x \in \Omega, \ t \in I := [0, T], \tag{4.5.1.1}$$

with initial condition

$$u(x, 0) = \hat{u}(x), \quad \forall x \in \Omega,$$

and boundary condition

$$u(x, t) = 0, \quad \forall x \in \partial\Omega, \ t \in I.$$

We assume that f satisfies

$$|f(u_1) - f(u_2)| \leq L_f ||u_1 - u_2||_0, \quad \forall u_1, \ u_2 \in \Lambda \subset \mathbb{R},$$

such that (4.5.1.1) possesses a unique solution $u \in \Lambda$. We refer [32] for the general description of the parabolic Fredholm equation. We begin our analysis with the weak form of (4.5.1.1),

$$(u_t, v) + ((u, v)) = (\int_\Omega f(u)dx, v), \ \forall v \in H_0^1(\Omega). \tag{4.5.1.2}$$

The corresponding P1 conforming finite element approximation for (4.5.1.2) is: Find $u_j \in V_j$ such that

$$(u_{j,t}, v) + ((u_j, v)) = (\int_\Omega f(u_j)dx, v), \quad \forall v \in V_j, \tag{4.5.1.3}$$

with $u_j(0) = R_j \hat{u}$, where $R_j : \ L^2(\Omega) \to V_j$ is defined by

$$((u - R_j u, v)) = 0, \quad \forall v \in V_j. \tag{4.5.1.4}$$

We know from Ciarlet [35] or Brenner and Scott [22] that

$$||u - R_j u||_0 \leq Ch_j^2, \tag{4.5.1.5}$$

$$|u - R_j u|_1 \leq Ch_j, \tag{4.5.1.6}$$

where the constant C is independent of h_j.

Theorem 4.5.1.1. *Let u and u_j be the solution of (4.5.1.2) and (4.5.1.3), respectively. Then we have*

$$||u - u_j||_0 \leq Ch_j^2, \tag{4.5.1.7}$$

$$|u - u_j|_1 \leq Ch_j, \tag{4.5.1.8}$$

where the constant C is independent of h_j.

Proof. Let

$$e := u - u_j = u - R_j u + R_j u - u_j := \rho + \theta.$$

Subtract (4.5.1.3) from (4.5.1.2),

$$(e_t, v) + ((e, v)) = (\int_\Omega [f(u) - f(u_j)]dx, v), \quad \forall v \in V_j. \tag{4.5.1.9}$$

Hence, we have the estimate

$$\begin{aligned}(\theta_t, v) + ((\theta. v)) &= (\int_\Omega [f(u) - f(R_j u)]dx. v) \\ &+ (\int_\Omega [f(R_j u) - f(u_j)]dx, v) - (\rho_t, v). \end{aligned} \quad (4.5.1.10)$$

for all $v \in V_j$, with $\theta(0) = \hat{u}_j - R_j\hat{u}$ in Ω, where

$$||\theta(0)||_0 \leq ||\hat{u}_j - \hat{u}||_0 + ||\hat{u} - R_j\hat{u}||_0 \leq Ch_j^2. \quad (4.5.1.11)$$

Taking $v = \theta$ in (4.5.1.10), we obtain

$$\begin{aligned}\frac{1}{2}\frac{d(||\theta||_0^2)}{dt} + |\theta|_1^2 &\leq L_f|\Omega| \cdot ||\theta||_0^2 + L_f|\Omega| \cdot ||\rho||_0||\theta||_0 + ||\rho_t||_0||\theta||_0 \\ &\leq (3L_f|\Omega| + 2)||\theta||_0 + 2L_f|\Omega|||\rho||_0 + 2||\rho_t||_0.\end{aligned}$$

Consequently, it holds that

$$\frac{d(||\theta||_0^2)}{dt} \leq 2(3L_f|\Omega| + 2)||\theta||_0^2 + 4L_f|\Omega|||\rho||_0^2 + 4||\rho_t||_0^2. \quad (4.5.1.12)$$

Integrate (4.5.1.12) with respect to t:

$$||\theta||_0^2 \leq ||\theta(0)||_0^2 + \int_0^t 2(3L_f|\Omega| + 2)||\theta||_0^2 ds + \int_0^t [4L_f|\Omega|||\rho||_0^2 + 4||\rho_t||_0^2]ds.$$

Gronwall's lemma leads to

$$||\theta||_0^2 \leq C\left(||\theta(0)||_0^2 + \int_0^t [||\rho||_0^2 + ||\rho_t||_0^2]ds\right).$$

Therefore, we have

$$||\theta||_0 \leq Ch_j^2, \quad (4.5.1.13)$$

where C is independent of h_j. Hence, we only need to derive (4.5.1.7).

To prove (4.5.1.7), choose $v = \theta_t$ in (4.5.1.10) and arrive at

$$\begin{aligned}
||\theta_t||_0^2 + \frac{1}{2}\frac{d(|\theta|_1^2)}{dt} &= (\int_\Omega [f(u) - f(R_j u)]dx, \theta_t) \\
&+ (\int_\Omega [f(R_j u) - f(u_j)]dx, \theta_t) - (\rho_t, \theta_t) \\
&\le || \int_\Omega [f(u) - f(R_j u)]dx||_0 \cdot ||\theta_t||_0 \\
&+ || \int_\Omega [f(R_j u) - f(u_j)]dx||_0 \cdot ||\theta_t||_0 + ||\rho_t||_0||\theta_t||_0 \\
&\le \frac{1}{2}||\theta_t||_0^2 + 2L_f^2|\Omega|^2||\rho||_0^2 + 2L_f^2|\Omega|^2||\theta||_0^2 + 2||\rho_t||_0^2.
\end{aligned}$$

After eliminating the first term on the right-hand side and integrating with respect to t, we obtain (in view of (4.5.1.13) and (4.5.1.11))

$$|\theta|_1 \le Ch_j,$$

where C is independent of h_j. Thus, we have verified Theorem 4.5.1.1.

4.5.2 Algorithm for parabolic Fredholm equation

Algorithm E: Let

$$u^0 = R_0 \hat{u}.$$

For $j = 2, \ldots, J$, solve

$$(u_t^j, v) + ((u^j, v)) = (\int_\Omega f(u^{j-1})dx, v)ds, \quad \forall v \in V_j, \tag{4.5.2.1}$$

for u^j.

Theorem 4.5.2.1. *Assume*

$$h_{j-1} \le \frac{C}{(2\sqrt{2})^j}. \tag{4.5.2.2}$$

Then

$$|u - u^J|_1 \le Ch_J,$$

where C is independent of h_J and the number of grids.

Proof. Let $e_j := u_j - u^j$. Subtracting (4.5.2.1) from (4.5.1.3) yields

$$(e_{j,t}, v) + ((e_j, v)) = (\int_\Omega [f(u_j) - f(u^{j-1})]dx, v)ds, \quad \forall v \in V_j, \tag{4.5.2.3}$$

Substitute $v = e_{j,t}$ into (4.5.2.3) to obtain

$$\begin{aligned} ||e_{j,t}||_0^2 + \frac{1}{2}\frac{d(|e_j|_1^2)}{dt} &\leq ||e_{j,t}||_0 || \int_\Omega [f(u_j) - f(u_{j-1})]dx||_0 \\ &+ ||e_{j,t}||_0 || \int_\Omega [f(u_{j-1}) - f(u^{j-1})]dx||_0 \\ &\leq \frac{1}{2}||e_{j,t}||_0^2 + L_f^2|\Omega|^2||u_j - u_{j-1}||_0^2 \\ &+ L_f^2|\Omega|^2||u_{j-1} - u^{j-1}||_0^2. \end{aligned} \tag{4.5.2.4}$$

Therefore, after eliminating the first term of right-hand side of (4.5.2.4) we find

$$\frac{d(|e_j|_1^2)}{dt} \leq Ch_{j-1}^4 + 2C|e_{j-1}|_1^2. \tag{4.5.2.5}$$

Integrating (4.5.2.5) leads to

$$|e_j|_1^2 \leq Ch_{j-1}^4 + 2C\int_0^t |e_{j-1}|_1^2 ds \leq Ch_{j-1}^4 + 2Ch_{j-2}^4 + \cdots + 2^{j-2}Ch_1^4 \leq Ch_j^2 8^j h_{j-1}^2.$$

In view of the assumption (4.5.2.2), we achieve

$$|e_j|_1 \leq Ch_j. \tag{4.5.2.6}$$

and hence, combining (4.5.2.2) with (4.5.1.8),

$$|u - u^J|_1 \leq |u - u_J|_1 + |u_J - u^J|_1 \leq Ch_J,$$

which completes the proof.

Chapter 5

The abstract cascading multigrid method in Besov spaces

It is important to mention that Bramble [19] contributed the general analysis of the V-cycle and the W-cycle in an abstract setting. Consider the FEM equation (5.2.1.2):

$$A_k y_k = Q_k g,$$

where A_k is symmetric positive definite. Let $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalue of A_k, respectively. The condition number of A_k is defined by

$$K(A_k) := \frac{\lambda_{\max}}{\lambda_{\min}}.$$

In view of the process of iteration, we know that the smaller the condition number $K(A_k)$ the more effective the iteration method is. Therefore, we define the preconditioning operator (or: preconditioner)

$$B_k : \ V_k \to V_k,$$

such that the condition number $K(B_k A_k)$ is as small as possible. Instead, we now solve the equation

$$B_k A_k y_k = B_k Q_k g,$$

by the iteration method. Indeed, this approach is more efficient. However the new problem arising is how to construct such a precondition operator B_k with as little computational cost as possible. Bramble [19] and the references mentioned there showed that the multigrid processes, the V-cycle and the W-cycle, are the ideal methods. We shall prove that the cascading multigrid process is also a simple and robust method. Bornemann and Deuflhard [15] provided the numerical examples to compare the cascading multigrid algorithm with the V-cycle algorithm. We analyze this by using the general framework with the assumptions abstracted from the FEM discretization of a given problem. In fact, the abstract multigrid framework provided us with a good way to describe the method for more complicated problems. To see this, we shall apply the cascading multigrid algorithms to the heat equation with mild regularity in Besov spaces [6] and to the equation discretized by the interior penalty discontinuous Galerkin method (see [52] and references therein). In comparison with the analysis in Shi and Xu [107], a distinctive feature of our method is the use of block Jacobi and symmetric Gauss-Seidel iteration as smoothers. We also extend these methods to VIDEs.

5.1 Notations and definitions

5.1.1 Bilinear form and induced norm

A bilinear form $A(\cdot,\cdot)$ in a Hilbert space V is called symmetric and elliptic, if it satisfies

1. (Symmetry) $A(u,v) = A(v,u), \quad \forall u,\ v \in V.$
2. (Continuity) There exists a positive constant C, such that,

$$A(u,v) \leq C||u||_V||v||_V, \quad \forall u,\ v \in V.$$

3. (Coercivity) There exists a positive constant c, such that,

$$A(u,u) \geq c||u||_V^2, \quad \forall u, \ v \in V.$$

Here $||\cdot||_V$ is the norm in V.

We find that the elliptic bilinear forms can define norms in V. Let the elliptic bilinear form $A(\cdot,\cdot)$ define the norm $|||\cdot|||$ by

$$|||\cdot||| := (A(\cdot,\cdot))^{1/2},$$

and let the elliptic mesh-dependent bilinear form $A_k(\cdot,\cdot)$ induce the norm $|||\cdot|||_k$ via

$$|||\cdot|||_k := (A_k(\cdot,\cdot))^{1/2}.$$

Define the time-dependent norm (see Section 5.2.2):

$$|||\cdot|||_\tau := (\tau^{-1}(\cdot,\cdot) + A(\cdot,\cdot))^{1/2},$$

and

$$|||\cdot|||_{k,\tau} := (\tau^{-1}(\cdot,\cdot) + A_k(\cdot,\cdot))^{1/2}.$$

Here $(\cdot,\cdot)$ is defined by

$$(u,v) := \int_\Omega uvdx, \quad \forall u, \ v \in V,$$

and the temporal mesh size τ is a given positive number.

5.1.2 Function spaces and their norms

We introduce the multi-index notation. A multi-index is defined as

$$\alpha := (\alpha_1, \ldots, \alpha_n), \quad \alpha_i \in \mathbb{N}_0.$$

The length of α is given by

$$|\alpha| := \sum_{i=1}^{n} \alpha_i.$$

For $\phi \in C^\infty$, we let

$$D^\alpha \phi := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \phi.$$

Given a vector $x := (x_1, \ldots, x_n)$, we define $x^\alpha := x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$.

Let k be nonnegative integer and let

$$w \in L^1_{loc}(\Omega) := \{w : \; w \in L^1(K), \; \forall \text{ compact } K \subset \text{interior } \Omega\}.$$

Suppose also that the weak derivatives $D^\alpha w$ exist for all $|\alpha| \leq k$. Define the Sobolev space norm

$$||w||_{H^k(\Omega)} := \left(\sum_{|\alpha| \leq k} ||D^\alpha w||^2_{L^2(\Omega)} \right)^{1/2}.$$

We define the Sobolev spaces via

$$H^k(\Omega) := \{w \in L^1_{loc}(\Omega) : \; ||w||_{H^k(\Omega)} < \infty\},$$

and

$$H^k_0(\Omega) := \{w \in H^k(\Omega) : \; w|_{\partial\Omega} = 0\}.$$

Let s be a nonnegative real number and let $[s]$ denote the integer part of s. Define the norm

$$||w||_{H^s(\Omega)} := \left(||w||^2_{H^{[s]}(\Omega)} + \sum_{|\alpha| = [s]} \int_\Omega \int_\Omega \frac{|w^\alpha(x) - w^\alpha(y)|^2}{|x-y|^{n+(s-[s])2}} dxdy \right)^{1/2}.$$

Then we define the fractional Sobolev spaces (Slobodeckeij space) via

$$H^s(\Omega) := \{w \in L^1_{loc}(\Omega) : \; ||w||_{H^s(\Omega)} < \infty\},$$

and

$$H^s_0(\Omega) := \{w \in H^s(\Omega) : \; w|_{\partial\Omega} = 0\}.$$

Remark 5.1.2.1. *We direct the readers to the book [1] for more general definitions and properties of Sobolev spaces.*

Kufner, John, and Fŭcik [79] and Triebel [114] discuss the method of defining fractional Sobolev spaces and Besov spaces by using the interpolation theory.

Let X and Y be two complex Banach spaces. Then we define, for $\theta \in (0,1)$, the interpolation space

$$[X, Y]_{s,\infty} := \{w \in Y : \sup_{t>0} t^{-2\theta} K(t,w)^2 < \infty\},$$

and the K-functional by

$$K(t,w) := \inf_{w_0 \in X} (||w_0||_X^2 + t^2 ||w - w_0||_Y^2)^{1/2}.$$

For an appropriate bounded domain $\Omega \subset \mathbb{R}^n$, we define the Besov space via

$$B^s(\Omega) := (H^{s_0}(\Omega), H^{s_1}(\Omega))_{\theta,\infty},$$

where $\theta \in (0,1)$, and s_0 and s_1 are nonnegative integers that satisfy $s_0 \neq s_1$, $s = (1-\theta)s_0 + \theta s_1$.

Remark 5.1.2.2. *If s is a fractional and positive number, then $H^s(\Omega) := B^s(\Omega)$ is the fractional Sobolev space.*

We now introduce the Besov space B_{01}^{-s} defined by the subspace interpolation of multilevel norms. The definition originates from [6].

Assume that

$$V_1 \subset V_2 \subset \cdots \subset V_k \subset \cdots$$

is a sequence of finite-dimensional subspaces of $H_0^1(\Omega)$ whose union is dense in $H_0^1(\Omega)$, and let Q_k denote the $L^2(\Omega)$ orthogonal projection onto V_k with $Q_0 := 0$ (see also

(5.2.1.1) of Section 5.2). Let

$$B^{-s}(\Omega) := [L^2(\Omega), H^{-1}(\Omega)]_{s,\infty}.$$

We introduce two norms, namely

$$||w||_{B_1^{-s}} := \sum_{k=1}^{\infty}(4^{k-1})^{-s/2}||(Q_k - Q_{k-1})w||_{L^2},$$

and

$$||w||^2_{B_{01}^{-s}} := ||w||^2_{B^{-s}} + ||w||^2_{B_1^{-s}}.$$

The space B_{01}^{-s} is then defined as the completion of $L^2(\Omega)$ with respect to the norm $||\cdot||_{B_{01}^{-s}}$.

5.2 The abstract cascading multigrid method

5.2.1 The cascading multigrid algorithm

Consider a finite-dimensional space V equipped with an inner product $(\cdot,\cdot)$ and a bilinear form $A(\cdot,\cdot)$, with corresponding norms $||\cdot||$ and $|||\cdot|||$. We further assume that $A(\cdot,\cdot)$ is elliptic (i.e., continuous and coercive) and symmetric. We let the subspaces V_k satisfy

$$V_0 \subset V_1 \subset \cdots \subset V_J \equiv V,$$

and we define the linear operator $A_k: \ V_k \longrightarrow V_k$ by

$$(A_k u, v) := A(u, v), \quad \text{for all } u, \ v \in V_k.$$

Obviously A_k is symmetric and positive definite. Furthermore, we define two projectors $P_k: \ V_J \longrightarrow V_k$ and $Q_k: \ V_J \longrightarrow V_k$ by

$$A(P_k u, v) := A(u, v),$$

and

$$(Q_k u, v) := (u, v), \tag{5.2.1.1}$$

for all $u \in V_J$ and $v \in V_k$. We also need to introduce a generic linear smoothing operator $R_k : V_k \to V_k$ with the assumption that $R_0 := A_0^{-1}$. For ease of analysis, we assume also that R_k is symmetric. To solve the equations for each level $k := 0, 1, \ldots, J$:

$$A_k y_k = Q_k g, \tag{5.2.1.2}$$

we shall now define the precondition operator $B \equiv B_J$ by the following cascading multigrid algorithm.

CMG Algorithm I:

0) $B_0 := A_0^{-1}$.

Define B_k implicitly in terms of B_{k-1}, for $k = 1, \ldots, J$:

1) For $\ell = 1, \ldots, m_k$, we set

$$y_k^\ell := y_k^{\ell-1} + R_k(Q_k g - A_k y_k^{\ell-1}).$$

Here $y_k^0 := B_{k-1} Q_{k-1} g$.

2) $B_k Q_k g := y_k^{m_k}$.

In order to analyze the CMG Algorithm I, we need to make the following four natural assumptions.

Assumption 5.2.1.1. *Let λ_k denote the maximum eigenvalue of A_k, i.e.,*

$$\lambda_k := \sup_{v \in V_k} \frac{(A_k v, v)}{(v, v)}. \tag{5.2.1.3}$$

For given $\alpha \in [0, 1]$, there exists C_α independent of k such that

$$(A_k^{1-\alpha}(I - P_{k-1})u, (I - P_{k-1})u) \le C_\alpha \lambda_k^{-\alpha} A(u, u), \quad \textit{for all } u \in V_k.$$

Assumption 5.2.1.2. *Let* $K_k := I - R_k A_k$. $R_{k,\omega} := \omega\lambda_k^{-1} I$ *and* $K_{k,\omega} := I - R_{k,\omega} A_k$. *There exists* $\omega \in (0, 1]$ *such that*

$$A(K_k v, K_k v) \leq A(K_{k,\omega/2} v, K_{k,\omega/2} v), \quad \text{for all } v \in V_k.$$

Hence it holds that

$$A(K_k^{m_k} v, K_k^{m_k} v) \leq A(K_{k,\omega/2}^{m_k} v, K_{k,\omega/2}^{m_k} v), \quad \text{for nonnegative number } m \text{ and for } v \in V_k.$$

Assumption 5.2.1.3. *We suppose that, for the given constant* $b > 0$,

$$c\frac{\lambda_J}{b^{J-k}} \leq \lambda_k \leq C\frac{\lambda_J}{b^{J-k}}.$$

Here and throughout the paper, c *and* C *denote generic positive constants which are independent of* k.

Assumption 5.2.1.4. *For* $\beta > 0$, *let*

$$m_k := \lfloor \beta^{J-k} m_J \rfloor,$$

where $\lfloor \cdot \rfloor$ *means the greatest integer function.*

Remark 5.2.1.1. *We see that Assumption 5.2.1.1 is reasonable, if we refer to the elliptic boundary value problems with less than full regularity (cf. [15] [105]).*

Remark 5.2.1.2. *Bramble [19] tells us that block Jacobi and (symmetric) Gauss-Seidel are the examples of* R_k *satisfying Assumption 5.2.1.2.*

We need also the following lemma about a property of $K_{k,\omega}$, which is necessary for the analysis of the cascading multigrid method.

Lemma 5.2.1.1. *The iterates of* $K_{k,\omega}^{m_k}$ *possess the following two properties:*

$$A(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \leq A(v, v), \tag{5.2.1.4}$$

and

$$A(K_{k,\omega}^{m_k}v, K_{k,\omega}^{m_k}v) \le C\frac{\lambda_k^{-\alpha}}{m_k^{\alpha}}(A^{1-\alpha}v, v), \tag{5.2.1.5}$$

for $\omega \in (0,1]$ *and* $\alpha \in (0,1]$.

Proof. See (pp. 164, Brenner and Scott [22]).

Theorem 5.2.1.1. *Let B_J be defined by CMG Algorithm I and assume that Assumptions {5.2.1.1, 5.2.1.2, 5.2.1.3, 5.2.1.4} hold. Then we have*

$$A((I - B_J A_J)u, u) \le \begin{cases} C \cdot \dfrac{1}{1-(b/\beta)^{\alpha/2}} \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u,u), & \text{for } \beta > b, \\ C \cdot J \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u,u), & \text{for } \beta = b, \end{cases}$$

for all $u \in V_J$.

Proof. We begin with the estimate:

$$\begin{aligned} y_J - B_J A_J y_J &= y_J - B_J Q_J g = y_J - y^{m_J} \\ &= K_J^{m_J}(y_J - y^0) = K_J^{m_J}(y_J - B_{J-1}Q_{J-1}g) \\ &= K_J^{m_J}(y_J - y_{J-1}) + K_J^{m_J}(y_{J-1} - B_{J-1}Q_{J-1}g) \\ &\quad \ldots\ldots \\ &= K_J^{m_J}(y_J - y_{J-1}) + K_J^{m_J}K_{J-1}^{m_{J-1}}(y_{J-1} - y_{J-2}) \\ &+ \ldots + K_J^{m_J}K_{J-1}^{m_{J-1}}\ldots K_1^{m_1}(y_1 - y_0). \end{aligned} \tag{5.2.1.6}$$

In view of (5.2.1.2), we obtain

$$y_k = A_k^{-1}Q_kQ_J^{-1}A_Jy_J = A_k^{-1}Q_kA_Jy_J, \quad \text{for } k = 0, \ldots, J-1. \tag{5.2.1.7}$$

Since Bramble [19] proved

$$Q_kA_J = A_kP_k, \quad \text{on } V_k,$$

we have

$$y_k = P_k y_J, \quad \text{for } k = 0, \ldots, J-1.$$

Bringing these equations into (5.2.1.6) leads to

$$I - B_J A_J = \sum_{k=1}^{J} \prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}} (P_k - P_{k-1}).$$

Consequently by using the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}
& A((I - B_J A_J)u, u) && (5.2.1.8)\\
= & \sum_{k=1}^{J} A\left(\prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}} (P_k - P_{k-1})u, u \right) \\
\leq & \sum_{k=1}^{J} [A\left(\prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}} (P_k - P_{k-1})u, \prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}} (P_k - P_{k-1})u \right)]^{1/2} \cdot [A(u,u)]^{1/2}.
\end{aligned}$$

In view of {Assumption 5.2.1.2, Lemma 5.2.1.1, Assumption 5.2.1.1}, we estimate

$$\begin{aligned}
& A\left(\prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}} (P_k - P_{k-1})u, \prod_{i=1}^{J-k+1} K_{J+1-i}^{m_{J+1-i}} (P_k - P_{k-1})u \right) \\
\leq & A\left(K_{k,\omega/2}^{m_k} (P_k - P_{k-1})u, K_{k,\omega/2}^{m_k} (P_k - P_{k-1})u \right) \\
\leq & C \frac{\lambda_k^{-\alpha}}{m_k^{\alpha}} (A^{1-\alpha}(P_k - P_{k-1})u, (P_k - P_{k-1})u) \\
\leq & C \frac{\lambda_k^{-\alpha}}{m_k^{\alpha}} \lambda_k^{-\alpha} A(u,u). && (5.2.1.9)
\end{aligned}$$

The estimates (5.2.1.8) and (5.2.1.9) lead to

$$A((I - B_J A_J)u, u) \leq C \left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}} \right) A(u,u). \qquad (5.2.1.10)$$

Combining (5.2.1.10), Assumption 5.2.1.3 and Assumption 5.2.1.4, we obtain our desired result.

Theorem 5.2.1.2. *Let $N_k := \dim(V_k)$ and assume $N_{k+1} \geq a N_k$. Then the computational cost of the CMG Algorithm I is given by*

$$\sum_{k=1}^{J} m_k N_k \leq \begin{cases} C \cdot \dfrac{1}{1-\beta/a} \cdot m_J N_J, & \text{for } \beta < a, \\ C \cdot J \cdot m_J N_J, & \text{for } \beta = a. \end{cases} \qquad (5.2.1.11)$$

Proof. $N_{k+1} \geq aN_k$ implies $N_k \leq a^{-(J-k)}N_J$. Recalling Assumption 5.2.1.4, we estimate

$$\sum_{k=1}^{J} m_k N_k \leq C \sum_{k=1}^{J} \left(\frac{\beta}{a}\right)^{J-k} \cdot m_J N_J.$$

Hence (5.2.1.11) is true.

Remark 5.2.1.3. *The parameter b in the expression of Theorem 5.2.1.1 solely depends on the original problem. It is well known that the finite element method for a partial differential equation of order n leads to $b = 2^n$. The parameter a in (5.2.1.11) (Theorem 5.2.1.2) depends only ion the dimension d of the domain, e.g.. $a = 2^d$. When $2^d = a > b$, that is, $d > \lfloor \log b \rfloor$, the condition number $K(I - B_J A_J)$ is uniformly bounded by $\mathcal{O}(\lambda_J^{-\alpha/2})$, and the amount of work is proportional to $\mathcal{O}(N_J)$. The cascading algorithm is henceforth called optimal. The algorithm will be called near-optimal if it satisfies the following corollary.*

Corollary 5.2.1.1 (Bornemann and Deuflhard (1996)). *In case of $b = a = 2^d$, we choose $\beta = 2^d$ and the number of iterations on level J as*

$$m_J = \lfloor m_* \cdot J^{d/\alpha} \rfloor.$$

Then we have the estimates

$$K(I - B_J A_J) \leq C\lambda_J^{-\alpha/2},$$

and

$$\sum_{k=1}^{J} m_k N_k \leq C m_* N_J (1 + \log N_J)^{1+\alpha/2}.$$

5.2.2 The method for the heat equation

Shi and Xu [107] analyzed the cascading multigrid method for the heat equation with Richardson and Conjugate Gradient iteration as smoothers. We shall solve the heat equation with mild regularity by the abstract cascading multigrid method with block Jacobi and symmetric Gauss-Seidel iteration as smoothers.

Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a open polygonal domain with the largest corner angle ϖ. We consider the heat equation:

$$u_t - \Delta u = f(x, t), \quad \text{on } \Omega \times [0, T], \tag{5.2.2.1}$$

with boundary condition

$$u(x, t) = 0, \quad \text{on } \partial\Omega \times [0, T],$$

and initial condition

$$u(x, 0) = u_0(x), \quad \text{on } \Omega,$$

where $f \in H^{-1+\alpha}(\Omega)$. The weak form of (5.2.2.1) is to find $u \in H_0^1(\Omega)$, with $u(x, 0) = u_0(x) \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$, such that

$$(u_t, v) + B(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega), \ t \in [0, T], \tag{5.2.2.2}$$

where the bilinear form B is

$$B(u, v) = \int_\Omega \nabla u \cdot \nabla v dx, \quad \forall u, \ v \in H_0^1(\Omega),$$

and

$$(f, v) = \int_\Omega f v dx.$$

We use the backward Euler scheme for the time-stepping. Let

$$I_\tau := \{t_n : \ 0 =: t_0 < t_1 < \cdots < t_N := T\},$$

be the mesh on I with

$$\tau_n := t_n - t_{n-1}, \ \tau := \max_{(n)} \tau_n.$$

Moreover, we set

$$w := u^n - u^{n-1}, \ (g, v) := (f, v) - B(u^{n-1}, v).$$

Then (5.2.2.2) leads to the weak form with time-stepping: Find $w \in H_0^1(\Omega)$ such that

$$A_\tau(w, v) = (g, v), \quad \forall v \in H_0^1(\Omega), \tag{5.2.2.3}$$

where

$$A_\tau(w, v) := \tau^{-1}(w, v) + B(w, v). \tag{5.2.2.4}$$

We conclude that (5.2.2.3) has a unique solution $w \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ which satisfies

$$||w||_{H^{1+\alpha}} \leq C||g||_{H^{-1+\alpha}}, \tag{5.2.2.5}$$

where $0 < \alpha < \alpha_0 = \frac{\pi}{\varpi}$ (see Ciarlet [35] and Johnson [76]).

Let $\mathcal{T}_k$ $(k := 0, 1, \ldots, J)$ be a sequence of quasi-uniform triangular partitions of Ω with mesh size $h_k = h_0 2^{-k}$. Let V_k denote the P1 conforming finite element space on $\mathcal{T}_k$. It is well known that

$$V_0 \subset V_1 \subset \cdots \subset V_J \equiv V \subset H_0^1(\Omega).$$

We derive the discrete form of (5.2.2.3): Find $w_k \in V_k$ such that

$$A_\tau(w_k, v) = (g, v), \quad \forall v \in V_k. \tag{5.2.2.6}$$

Define

$$(A_k w_k, v) := A_\tau(w_k, v), \quad \forall w_k, \ v \in V_k.$$

Then (5.2.2.6) can be expressed by

$$A_k w_k = g_k, \tag{5.2.2.7}$$

where $g_k \in V_k$, $(g_k, v) := (g, v)$, $\forall v \in V_k$. It is well-known that

$$|||w - w_k|||_\tau \leq C h_k^\alpha (1 + \tau^{-1} h_k)^{1/2} ||g||_{H^{-1+\alpha}}, \tag{5.2.2.8}$$

where

$$|||\cdot|||_\tau := \left(\tau^{-1}(\cdot,\cdot) + B(\cdot,\cdot)\right)^{1/2}.$$

We know from the definitions (5.2.1.3) and (5.2.2.4) that

$$\lambda_k = \mathcal{O}(h_k^{-2}) + \tau^{-1}.$$

Since [10] observed that some commonly used iterative methods, like Richardson iteration, can already guarantee good convergence for $\tau \leq \lambda_J^{-1}$, we therefore only consider the case

$$\lambda_J^{-1} \leq \tau \leq \frac{\lambda_J^{-1}}{\gamma_0}, \quad \text{for some } \gamma_0 \in (0,1). \tag{5.2.2.9}$$

Now we shall verify that Assumptions {(5.2.1.1), (5.2.1.2), (5.2.1.3), (5.2.1.4)} hold true in this case. Let

$$A := A_\tau.$$

From [19] and [113], we conclude that Assumption 5.2.1.1 is satisfied. Assumption 5.2.1.3 can be easily written as

$$c\left(\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right) \leq \lambda_k \leq C\left(\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right). \tag{5.2.2.10}$$

Assumption 5.2.1.2 and {(5.3.1.2), (5.3.1.3)} concern only R_k. Therefore they remain true as discussed in Section 5.2.1. Thus, recalling (5.2.1.10), we obtain

$$\begin{aligned} & A((I - B_J A_J)w, w) \qquad (5.2.2.11) \\ \leq\ & C\left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}}\right) A(w,w) \\ \leq\ & C\left(\frac{1}{1+\gamma_0}\right)^\alpha \frac{h_J^\alpha}{m_J^{\alpha/2}} \sum_{k=1}^{J}\left(\frac{2^\alpha}{\beta^{\alpha/2}}\right)^k A(w,w). \end{aligned}$$

So (5.2.2.11) yields the following theorem. The last estimate in (5.2.2.11) reveals that the attainable (optimal or nearly optimal) order depends on the value of β; for $\beta = 4$ the estimate will contain the factor J, i.e., the number of the grid points. Compare also Bornemann and Deuflhard [15].

Theorem 5.2.2.1. *Under Assumption 5.2.1.4, we have the following convergence estimate of CMG Algorithm I for (5.2.2.7):*

$$A((I - B_J A_J)w, w) \leq \begin{cases} C(\alpha) \cdot \dfrac{1}{1 - \left(\dfrac{2}{\beta^{1/2}}\right)^{\alpha}} \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \text{for } \beta > 4, \\ C(\alpha) \cdot J \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \text{for } \beta = 4, \end{cases}$$

with $C(\alpha) := C(1/(1 + \gamma_0))^{\alpha}$. These estimates hold for all $w \in V_J$.

In view of

$$N_{k+1} \geq 2^d N_k,$$

we conclude that Theorem 5.2.1.2 remains true with $a = 2^d$. Clearly, when $d = 2$, CMG Algorithm I for (5.2.2.7) is near optimal; when $d = 3$, the algorithm is optimal.

For a detailed discussion on how to choose β so that the cascading multigrid method has optimal accuracy and complexity, we refer to Shi and Xu [107] or Bornamann and Deuflhard [15].

5.2.3 The cascading multigrid method in Besov spaces

For $s \in (0, 1)$, define the interpolation space

$$[X, Y]_{s,\infty} := \{w \in Y : \sup_{t>0} t^{-2s} K(t, w)^2 < \infty\},$$

and the K-functional by

$$K(t, w) := \inf_{w_0 \in X} (||w_0||_X^2 + t^2 ||w - w_0||_Y^2)^{1/2}.$$

Recall the equation (5.2.2.3): Find $w \in H_0^1(\Omega)$ such that

$$A_\tau(w, v) = (g, v), \quad \forall v \in H_0^1(\Omega), \tag{5.2.3.1}$$

where

$$A_\tau(w, v) := \tau^{-1}(w, v) + B(w, v).$$

Recall (5.2.2.5): for $0 < \alpha < \alpha_0 = \frac{\pi}{\varpi}$, we have the regularity

$$||w||_{H^{1+\alpha}} \leq C||g||_{H^{-1+\alpha}}, \quad \forall g \in H^{-1+\alpha_0}(\Omega).$$

For the critical case $\alpha = \alpha_0$, Bacuta, Bramble and Xu [6] proved the estimate

$$||w||_{B^{1+\alpha_0}(\Omega)} \leq C||g||_{B_{01}^{-1+\alpha_0}(\Omega)}, \quad \forall g \in B_{01}^{-1+\alpha_0}(\Omega), \tag{5.2.3.2}$$

where $B^{1+\alpha_0}(\Omega)$ is a standard Besov space and $B_{01}^{-1+\alpha_0}(\Omega)$ was defined in [6]. Correspondingly, the convergence estimate of the P1 finite element method for (5.2.3.1) is

$$\begin{aligned} |||w - w_k|||_{k,\tau} &\leq Ch_k^{\alpha_0}(1 + \tau^{-1}h_k)^{1/2}||w||_{[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-\alpha_0,\infty}} \\ &\leq Ch_k^{\alpha_0}(1 + \tau^{-1}h_k)^{1/2}||g||_{B_{01}^{-1+\alpha_0}(\Omega)}. \end{aligned} \tag{5.2.3.3}$$

We begin our analysis with an analogue of equation (5.2.1.9). Triebel [114, pp. 59] proved the interpolation property

$$H^1(\Omega) = [H^1(\Omega), H^1(\Omega)]_{\theta,\infty}.$$

Hence, we have

$$\begin{aligned} & A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w, K_{k,w/2}^{m_k}(P_k - P_{k-1})w\right) \\ \leq\ & C\sup_{t>0} t^{-2(1-\alpha_0)} \inf_{w_0\in V_J} [A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0, K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0\right) \\ +\ & t^2 A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0), K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0)\right)]. \end{aligned} \tag{5.2.3.4}$$

We use Lemma 5.2.1.1 and Assumption 5.2.1.1 with $\alpha = 0$ and $\alpha = 1$ to get the equations

$$\begin{aligned} & A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0, K_{k,w/2}^{m_k}(P_k - P_{k-1})w_0\right) \\ \leq \; & C\frac{\lambda_k^{-1}}{m_k}\left((P_k - P_{k-1})w_0, (P_k - P_{k-1})w_0\right) \leq C\frac{\lambda_k^{-2}}{m_k}A(w_0, w_0), \quad (5.2.3.5) \end{aligned}$$

and

$$A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0), K_{k,w/2}^{m_k}(P_k - P_{k-1})(w - w_0)\right) \leq CA(w - w_0, w - w_0). \quad (5.2.3.6)$$

Hence $\{(5.2.3.4), (5.2.3.5), (5.2.3.6)\}$ lead to

$$\begin{aligned} & A\left(K_{k,w/2}^{m_k}(P_k - P_{k-1})w, K_{k,w/2}^{m_k}(P_k - P_{k-1})w\right) \\ \leq \; & C\frac{\lambda_k^{-2\alpha_0}}{m_k^{\alpha_0}}\sup_{t>0}(m_k^{1/2}\lambda_k t)^{-2(1-\alpha_0)}\inf_{w_0\in V_J}\left[A(w_0, w_0) + (m_k^{1/2}\lambda_k t)^2 A(w - w_0, w - w_0)\right] \\ \leq \; & C\frac{\lambda_k^{-2\alpha_0}}{m_k^{\alpha_0/2}}A(w, w). \quad (5.2.3.7) \end{aligned}$$

Therefore we obtain

$$A((I - B_J A_J)w, w) \leq C\left(\sum_{k=1}^{J}\frac{\lambda_k^{-\alpha_0}}{m_k^{\alpha_0/2}}\right)A(w, w).$$

This suggests the following theorem:

Theorem 5.2.3.1. *Under Assumption 5.2.1.4, we have the convergence estimate of the CMG Algorithm I for (5.2.3.1), namely*

$$A((I - B_J A_J)w, w) \leq \begin{cases} C \cdot \dfrac{1}{1 - \left(\dfrac{2}{\beta^{1/2}}\right)^{\alpha_0}} \cdot \dfrac{h_J^{\alpha_0}}{m_J^{\alpha_0/2}}A(w, w), & \text{for } \beta > 4, \\ C \cdot J \cdot \dfrac{h_J^{\alpha_0}}{m_J^{\alpha_0/2}}A(w, w), & \text{for } \beta = 4, \end{cases}$$

for all $w \in V_J$.

5.2.4 The method for VIDEs

We consider

$$u_t - \triangle u + \int_0^t k(t-s)\mathcal{B}u(x,s)ds = f(x,t), \quad x \in \Omega, \ t \in I := [0,T], \qquad (5.2.4.1)$$

with initial condition

$$u(x,0) = u_0(x), \quad x \in \Omega,$$

and boundary condition

$$u(x,t) = 0, \quad x \in \partial\Omega, \ t \in I,$$

where $\Omega \subset \mathbb{R}^d$ $(d = 2,3)$ is a bounded convex polygonal domain. $\mathcal{B}$ is an elliptic differential operator of order up to two. We can easily follow the idea of Chapter 4 to analyze the cascading multigrid method for (5.2.4.1) with less than second-order $\mathcal{B}$. So it is interesting to analyze the case with dominant memory term (i.e., the operator $\mathcal{B}$ is second order). For ease of exposition, we suppose $\mathcal{B} := -\triangle$. In most cases of application, k is nonnegative. We want to use the trapezoidal rule (5.2.4.3) for the memory term, hence we assume $k \in C^1(I)$. If k is weakly singular, we can use the left-rectangular rule (see [34]).

The weak form of (5.2.4.1) is to find $u \in H_0^1(\Omega)$, with $u(x,0) = u_0(x) \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$, such that

$$(u_t, v) + B(u,v) + \int_0^t k(t-s)B(u(x,s),v)ds = (f,v), \ \ \forall v \in H_0^1(\Omega), \ \ t \in I, \ \ (5.2.4.2)$$

where the bilinear form B is as before:

$$B(u,v) = \int_\Omega \nabla u \cdot \nabla v dx, \quad \forall u, \ v \in H_0^1(\Omega),$$

and also

$$(f,v) = \int_\Omega f v dx.$$

We use the backward Euler scheme for time-stepping of (5.2.4.2) and the trapezoidal rule,

$$\int_0^{t_n} p(s)ds \approx \frac{1}{2}\tau p(0) + \sum_{j=1}^{n-1} \tau p(t_j) + \frac{1}{2}\tau p(t_n), \tag{5.2.4.3}$$

for discretizing the memory term.

Let $k_j := k(t_n - t_j)$ $(j = 0, \ldots, n)$. Then we get that: Find $u^n - u^{n-1} =: w \in H_0^{1+\alpha}(\Omega)$, such that

$$A_\tau(w, v) = (g, v), \quad \forall v \in H_0^1(\Omega), \tag{5.2.4.4}$$

where

$$A_\tau(w, v) := \tau^{-1}(w, v) + (1 + \frac{1}{2}\tau k_n)B(w, v),$$

and

$$(g, v) := (f, v) - (1 + \frac{1}{2}\tau k_n)B(u^{n-1}, v) - \frac{1}{2}\tau k_0 B(u_0, v) - \sum_{j=1}^{n-1} \tau k_j B(u^j, v).$$

The same discussion as that in Section 5.2.2 leads to the P1 conforming finite element approximation to (5.2.4.4): Find $w_k \in V_k \subset H_0^1(\Omega)$, such that

$$A_\tau(w_k, v) = (g, v), \quad \forall v \in V_k, \tag{5.2.4.5}$$

and its error estimate is

$$|||w - w_k|||_\tau \leq C h_k^\alpha (1 + \tau^{-1} k_0 h_k)^{1/2} ||g||_{H^{-1+\alpha}}. \tag{5.2.4.6}$$

We also define

$$(A_k w_k, v) := A_\tau(w_k, v), \quad \forall w_k, \ v \in V_k. \tag{5.2.4.7}$$

Then (5.2.4.5) can be written as

$$A_k w_k = g_k, \tag{5.2.4.8}$$

where

$$g_k \in V_k, \ (g_k, v) := (g, v), \ \forall v \in V_k.$$

We see that

$$\lambda_k = C(1 + \frac{1}{2}\tau k_0) h_k^{-2} + \tau^{-1}.$$

Hence, we have

$$c\left(\frac{(1+\frac{1}{2}\tau k_0)\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right) \le \lambda_k \le C\left(\frac{(1+\frac{1}{2}\tau k_0)\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right). \tag{5.2.4.9}$$

Assumption 5.2.1.2 and {(5.3.1.2), (5.3.1.3)} concern only R_k. They henceforth remain true as discussed in Section 5.2.1. Thus, recalling (5.2.1.10), we obtain

$$\begin{aligned} A((I - B_J A_J)w, w) &\le C\left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}}\right) A(w, w) \qquad (5.2.4.10) \\ &\le C\frac{h_J^{\alpha}}{m_J^{\alpha/2}} \sum_{k=1}^{J} \left(\frac{2^{\alpha}}{\beta^{\alpha/2}}\right)^k A(w, w). \end{aligned}$$

We used $a + b \ge 2\sqrt{ab}$ in the calculation from (5.2.4.9) to (5.2.4.10).

So (5.2.4.10) yields the following theorem.

Theorem 5.2.4.1. *Under Assumption 5.2.1.4, we have the following convergence estimate of the CMG Algorithm I for (5.2.4.8):*

$$\begin{aligned} &A((I - B_J A_J)w, w) \\ \le &\begin{cases} C \cdot \dfrac{1}{1 - \left(\dfrac{2}{\beta^{1/2}}\right)^{\alpha}} \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \textit{for } \beta > 4, \\ C \cdot J \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \textit{for } \beta = 4, \end{cases} \end{aligned}$$

for all $w \in V_J$.

Remark 5.2.4.1. *We notice that compared with Section 5.2.2 and Section 5.3.2, the analysis of this section does not need the assumption (5.2.2.9).*

5.3 Abstract cascading multigrid method with non-nested spaces and varying inner products and bilinear forms

In the previous sections we assumed that $V_0 \subset V_1 \subset \cdots \subset V_J \equiv V$, and that $(\cdot, \cdot)$ and $A(\cdot, \cdot)$ are defined on all V_k. In this section, we allow the spaces V_k to be not necessarily nested and define a symmetric elliptic bilinear form $A_k(\cdot, \cdot)$ and an inner product $(\cdot, \cdot)_k$ corresponding to $0 \leq k \leq J$.

5.3.1 The cascading multigrid algorithm

Consider the following setting:

(1) Let $V_0, \ldots, V_J$ be J finite dimensional spaces with $V_J \equiv V$, which are not necessarily nested.

(2) Define J linear operators

$$I_k : \; V_{k-1} \longrightarrow V_k, \quad k = 1, \ldots, J,$$

which connect the spaces.

(3) Let $(\cdot, \cdot)_k$ be an inner product on $V_k \times V_k$, with induced norm $|| \cdot ||_k$, and let $A_k(\cdot, \cdot)$ be symmetric and elliptic bilinear form on $V_k \times V_k$ with $A_J(\cdot, \cdot) \equiv A(\cdot, \cdot)$ and induced norm $||| \cdot |||_k$.

(4) Define $A_k : \; V_k \longrightarrow V_k$ by

$$(A_k u, v)_k := A_k(u, v), \quad \text{for all } v \in V_k.$$

(5) Define $P_{k-1} : \; V_k \longrightarrow V_{k-1}$ and $Q_{k-1} : \; V_k \longrightarrow V_{k-1}$ by

$$A_{k-1}(P_{k-1} u, v) := A_k(u, I_k v),$$

and

$$(Q_{k-1}u, v)_{k-1} := (u, I_k v)_k,$$

for all $v \in V_{k-1}$.

(6) Let R_k be a linear symmetric operator $R_k : V_k \to V_k$ with $R_0 := A_0^{-1}$.

We shall solve the equation (5.2.1.2) on each level $k := 0, 1, \ldots, J$ by the cascading multigrid algorithm.

Define the precondition operator $B \equiv B_J$ by the

CMG Algorithm II:

0) $B_0 := A_0^{-1}$;

Define B_k implicitly in terms of B_{k-1}, for $k = 1, \ldots, J$:

1) For $\ell = 1, \ldots, m_k$, we set

$$y_k^\ell := y_k^{\ell-1} + R_k(Q_k g - A_k y_k^{\ell-1}).$$

Here $y_k^0 := I_k B_{k-1} Q_{k-1} g$.

3) $B_k Q_k g := y_k^{m_k}$.

Before we analyze the CMG Algorithm II, we add the following three natural assumptions.

Assumption 5.3.1.1. *We suppose*

$$A_k(I_k u, I_k u) \leq A_{k-1}(u, u), \quad \text{for all } u \in V_{k-1}. \tag{5.3.1.1}$$

Using the Cauchy-Schwartz inequality, we can easily prove that (5.3.1.1) holds true if and only if

$$A_{k-1}(P_{k-1}u, P_{k-1}u) \leq A_k(u, u), \quad \text{for all } u \in V_k.$$

Assumption 5.3.1.2. *Let λ_k denote the maximum eigenvalue of A_k, i.e.,*

$$\lambda_k := \sup_{v \in V_k} \frac{(A_k v, v)_k}{(v, v)_k}.$$

For given $\alpha \in (0, 1]$, there exists a constant C_α, independent of k, such that

$$(A_k^{1-\alpha}(I - I_k P_{k-1})u, (I - I_k P_{k-1})u)_k \le C_\alpha \lambda_k^{-\alpha} A_k(u, u), \quad \textit{for all } u \in V_k$$

(compare with Assumption 5.2.1.1, $(\cdot, \cdot) \equiv (\cdot, \cdot)_k$).

Assumption 5.3.1.3. *Let $K_k := I - R_k A_k$, $R_{k,\omega} := \omega \lambda_k^{-1} I$ and $K_{k,\omega} := I - R_{k,\omega} A_k$. There exists $\omega \in (0, 1]$ such that*

$$A_k(K_k v, K_k v) \le A_k(K_{k,\omega/2} v, K_{k,\omega/2} v), \quad \textit{for all } v \in V_k.$$

Hence it holds that

$$A_k(K_k^{m_k} v, K_k^{m_k} v) \le A_k(K_{k,\omega/2}^{m_k} v, K_{k,\omega/2}^{m_k} v), \ \textit{for nonnegative number } m \textit{ and for } v \in V_k$$

(compare with Assumption 5.2.1.2).

By using the same argument as in the proof of Lemma 5.2.1.1, we can prove the following lemma.

Lemma 5.3.1.1. *$K_{k,\omega}^{m_k}$ satisfies the following two properties:*

$$A_k(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \le A_k(v, v), \tag{5.3.1.2}$$

and

$$A_k(K_{k,\omega}^{m_k} v, K_{k,\omega}^{m_k} v) \le C \frac{\lambda_k^{-\alpha}}{m_k^\alpha} (A_k^{1-\alpha} v, v)_k, \tag{5.3.1.3}$$

for $\omega \in (0, 1]$ and $\alpha \in (0, 1]$.

Now we are ready to present our main results.

Theorem 5.3.1.1. *Let B_J be defined by the CMG Algorithm II, and let Assumptions* $\{5.3.1.1, 5.3.1.2, 5.3.1.3, 5.2.1.3, 5.2.1.4\}$ *hold. Then we have*

$$A((I - B_J A_J)u, u) \leq \begin{cases} C \cdot \dfrac{1}{1-(b/\beta)^{\alpha/2}} \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u,u), & \text{for } \beta > b, \\ C \cdot J \cdot \dfrac{\lambda_J^{-\alpha/2}}{m_J^{\alpha/2}} A(u,u), & \text{for } \beta = b, \end{cases}$$

for all $u \in V_J$.

Proof. We first estimate:

$$\begin{aligned} y_J - B_J A_J y_J &= y_J - B_J Q_J g = y_J - y^{m_J} \qquad (5.3.1.4) \\ &= K_J^{m_J}(y_J - y^0) = K_J^{m_J}(y_J - I_J B_{J-1} Q_{J-1} g) \\ &= K_J^{m_J}(y_J - I_J y_{J-1}) + K_J^{m_J}(I_J y_{J-1} - I_J B_{J-1} Q_{J-1} g) \\ &\quad \ldots\ldots \\ &= K_J^{m_J}(y_J - I_J y_{J-1}) + (K_J^{m_J} I_J) K_{J-1}^{m_{J-1}}(y_{J-1} - I_{J-1} y_{J-2}) \\ &+ \ldots + (K_J^{m_J} I_J)(K_{J-1}^{m_{J-1}} I_{J-1}) \ldots (K_2^{m_2} I_2) K_1^{m_1}(y_1 - I_1 y_0). \end{aligned}$$

From (5.2.1.2):

$$A_k y_k = Q_k g,$$

we derive

$$(A_k y_k, I_k v)_k = (Q_k g, I_k v)_k, \quad \forall v \in V_{k-1}.$$

Hence,

$$A_k(y_k, I_k v) = (Q_k g, I_k v)_k, \quad \forall v \in V_{k-1}.$$

In view of the definitions of P_k and Q_k, we obtain

$$A_{k-1}(P_{k-1} y_k, v) = (Q_{k-1} g, v)_{k-1}, \quad \forall v \in V_{k-1},$$

which leads to

$$A_{k-1}P_{k-1}y_k = Q_{k-1}g.$$

Therefore, we have

$$y_{k-1} = P_{k-1}y_k, \quad \text{for } k = 0, \dots, J-1. \tag{5.3.1.5}$$

Combining (5.3.1.5) and (5.3.1.4) we find

$$I - B_J A_J = K_J^{m_J}(I - I_J P_{J-1}) + \sum_{k=1}^{J-1}\prod_{i=1}^{J-k}(K_{J+1-i}^{m_{J+1-i}} I_{J+1-i}) \cdot K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i}.$$

By the Cauchy-Schwarz inequality and Lemma 5.3.1.1,

$$\begin{aligned} & A((I - B_J A_J)u, u) \\ \leq\ & [A(K_J^{m_J}(I - I_J P_{J-1})u, K_J^{m_J}(I - I_J P_{J-1})u)]^{1/2} \cdot [A(u,u)]^{1/2} \\ +\ & \sum_{k=1}^{J-1}[A_k\left((K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u, (K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u\right)]^{1/2} \\ \cdot\ & [A(u,u)]^{1/2}. \end{aligned} \tag{5.3.1.6}$$

From Assumption 5.3.1.3, Lemma 5.3.1.1, and Assumption 5.3.1.2, we have

$$\begin{aligned} & A\left(K_J^{m_J}(I - I_J P_{J-1})u, K_J^{m_J}(I - I_J P_{J-1})u\right) \\ \leq\ & C\frac{\lambda_J^{-\alpha}}{m_J^\alpha}\left(A^{1-\alpha}(I - I_J P_{J-1})u, (I - I_J P_{J-1})u\right) \\ \leq\ & C\frac{\lambda_J^{-\alpha}}{m_J^\alpha}A(u,u), \end{aligned} \tag{5.3.1.7}$$

and in addition, using Assumption 5.3.1.1,

$$\begin{aligned} & A_k\left((K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u, (K_k^{m_k} \cdot (I - I_k P_{k-1}) \cdot \prod_{i=1}^{J-k} P_{J-i})u\right) \\ \leq\ & C\frac{\lambda_k^{-\alpha}}{m_k^\alpha}\left(A_k^{1-\alpha}(I - I_k P_{k-1})\prod_{i=1}^{J-k} P_{J-i}u, (I - I_k P_{k-1})\prod_{i=1}^{J-k} P_{J-i}u\right)_k \\ \leq\ & C\frac{\lambda_k^{-\alpha}}{m_k^\alpha}A_k\left(\prod_{i=1}^{J-k} P_{J-i}u, \prod_{i=1}^{J-k} P_{J-i}u\right) \leq C\frac{\lambda_k^{-\alpha}}{m_k^\alpha}A(u,u). \end{aligned} \tag{5.3.1.8}$$

Hence (5.3.1.6), (5.3.1.7) and (5.3.1.8) lead to

$$A((I - B_J A_J)u, u) \le C \left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}} \right) A(u, u). \tag{5.3.1.9}$$

Combining (5.3.1.9) and Assumption 5.2.1.3 and Assumption 5.2.1.4, we obtain our desired result.

The computational cost estimate theorem is the same as Theorem 5.2.1.2.

5.3.2 Application to interior penalty discontinuous Galerkin method

The V-cycle algorithm for the interior penalty discontinuous Galerkin method was presented in the paper by Gopalakrishnan and Kanschat [52], which was based on Arnold [3]. In this section, we analyze the CMG Algorithm II for the heat equation (5.2.2.1) with the interior penalty discontinuous Galerkin discretization. We will essentially use the notations in Section 5.3.1.

Let $\mathcal{T}_k$ $(k := 0, 1, \dots, J)$ be a quasi-uniform triangular partition of Ω with the mesh size $h_k = h_0 2^{-k}$. We define the multilevel spaces

$$V_0 \subset V_1 \subset \cdots \subset V_J \equiv V,$$

by

$$V_k := \{v : \; v|_T \in \mathcal{P}^{(m)}(T), \; \forall T \in \mathcal{T}_k\},$$

where $\mathcal{P}^{(m)}(T)$ denotes the polynomial with degree not exceeding m on T.

To describe the interior penalty discontinuous Galerkin method, we need the spaces

$$H_0^1(\mathcal{T}_k) := \{v \in L^2(\Omega) : \; v|_T \in H^1(T) \text{ and } v|_{\partial\Omega \cap T} = 0, \; \forall T \in \mathcal{T}_k\}.$$

Let $\mathcal{E}_k$ denote the set of edges of the triangulation $\mathcal{T}_k$. If $e \in \mathcal{E}_k$ is an interior edge, we denote by n_e one of the two unit normal vectors at e, and we define the jumps and averages of the normal derivatives (for $x \in e$) of $v \in H_0^1(\mathcal{T}_k)$ by

$$[v]_e(x) := \lim_{\delta \to 0+} [v(x - \delta n_e) - v(x + \delta n_e)],$$

and

$$\langle \partial_n v \rangle_e(x) := \frac{1}{2} \lim_{\delta \to 0+} [n_e \cdot \nabla v(x - \delta n_e) + n_e \cdot \nabla v(x + \delta n_e)].$$

If $e \subseteq \partial\Omega$, we fix n_e to be the outward normal vector and let

$$[v]_e(x) := \lim_{\delta \to 0+} v(x - \delta n_e) \text{ and } \langle \partial_n v \rangle_e := \lim_{\delta \to 0+} n_e \cdot \nabla v(x - \delta n_e).$$

Define $B_k(\cdot, \cdot)$ on $H_0^1(\mathcal{T}_k) \times H_0^1(\mathcal{T}_k)$ by

$$\begin{aligned} B_k(u, v) \quad &:= \quad \sum_{T \in \mathcal{T}_k} (\nabla u, \nabla v)_T \\ &\quad + \sum_{e \in \mathcal{E}_k} \left(\frac{\sigma}{\ell_e} ([u], [v])_e - (\langle \partial_n u \rangle, [v])_e - ([u], \langle \partial_n v \rangle)_e \right). \end{aligned} \tag{5.3.2.1}$$

Here ℓ_e denotes the length of the edge e and σ is a positive parameter to be chosen later.

The weak form of (5.2.2.1) is defined by: Find $u \in H_0^1(\mathcal{T}_k)$, with $u(x, 0) = u_0(x) \in H^{1+\alpha}(\mathcal{T}_k) \cap H_0^1(\mathcal{T}_k)$, such that

$$(u_t, v) + B_k(u, v) = (f, v), \quad \forall v \in H_0^1(\mathcal{T}_k), \; t \in [0, T]. \tag{5.3.2.2}$$

As in Section 5.2.2, we use the backward Euler scheme for the time-stepping of (5.3.2.2). Then we derive the weak form with time-stepping: To find $w \in H_0^{1+\alpha}(\mathcal{T}_k)$ such that

$$A_{\tau,k}(w, v) := (g, v), \quad \forall v \in H_0^1(\mathcal{T}_k), \tag{5.3.2.3}$$

where

$$A_{\tau,k}(w,v) := \tau^{-1}(w,v) + B_k(w,v),$$

and

$$(g,v) := (f,v) - B_k(u^{n-1},v).$$

The interior penalty discontinuous Galerkin approximation to (5.3.2.3):

Find $w_k \in V_k$, such that

$$A_{\tau,k}(w_k,v) = (g,v), \quad \forall v \in V_k. \tag{5.3.2.4}$$

Define

$$(A_k w_k, v) := A_k(w_k, v) := A_{\tau,k}(w_k,v), \quad \forall w_k, \ v \in V_k. \tag{5.3.2.5}$$

Then (5.3.2.4) can be expressed by

$$A_k w_k = g_k, \tag{5.3.2.6}$$

where

$$g_k \in V_k, \ (g_k, v) := (g,v), \ \forall v \in V_k.$$

From [3], we can easily obtain that

$$|||w - w_k|||_{k,\tau} \le C h_k^{\alpha}(1 + \tau^{-1}h_k)^{1/2}||g||_{H^{-1+\alpha}}. \tag{5.3.2.7}$$

From [52], we know that

$$c(\sigma+1)h_k^{-2} + \tau^{-1} \le \lambda_k \le C(\sigma+1)h_k^{-2} + \tau^{-1}.$$

Hence we have

$$c\left((\sigma+1)\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right) \le \lambda_k \le C\left((\sigma+1)\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right). \tag{5.3.2.8}$$

In [52] it was also proved that Assumption 5.3.1.1 and Assumption 5.3.1.2 hold true. Assumption 5.3.1.3 and Lemma 5.3.1.1 are guaranteed, since they only concern the smoother operator R_k. We assume τ satisfies (5.2.2.9). So by recalling (5.3.1.9), we derive that

$$\begin{aligned}
&A((I - B_J A_J)w, w) \\
\leq\ & C \left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}} \right) A(w, w) \\
\leq\ & C \left(\frac{1}{1 + \sigma + \gamma_0} \right)^{\alpha} \frac{h_J^{\alpha}}{m_J^{\alpha/2}} \sum_{k=1}^{J} \left(\frac{2^{\alpha}}{\beta^{\alpha/2}} \right)^{k} A(w, w).
\end{aligned}$$

The following theorem is therefore established.

Theorem 5.3.2.1. *Under Assumption 5.2.1.4, we have the convergence estimate of CMG Algorithm II for (5.3.2.6):*

$$A((I - B_J A_J)w, w) \leq \begin{cases} C \cdot \dfrac{1}{1 - \left(\dfrac{2}{\beta^{1/2}} \right)^{\alpha}} \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \text{for } \beta > 4, \\ C \cdot J \cdot \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \text{for } \beta = 4. \end{cases}$$

for all $w \in V_J$.

The computational cost estimate is that of Theorem 5.2.1.2 with $a = 2^d$.

Remark 5.3.2.1. *The foregoing discussion reveals that the abstract setting of the cascading multigrid method provides a more feasible way to establish the convergence theorem of the method for the time-dependent problems with mild regularity, and for problems that are discretized by other new FEM methods. We believe it is also possible to extend this abstract framework to the mesh-free method described in, e.g., [14, 85].*

5.3.3 Extension to VIDEs

In this section, we shall extend the CMG Algorithm II to (5.2.4.1) with the interior penalty discontinuous Galerkin approximation. We start from the interior penalty discontinuous Galerkin weak form of (5.2.4.1).

Find $u \in H_0^1(\mathcal{T}_k)$ with $u(x,0) = u_0(x) \in H^{1+\alpha}(\mathcal{T}_k) \cap H_0^1(\mathcal{T}_k)$, such that

$$(u_t, v) + B_k(u,v) + \int_0^t B_k(u(x,s),v)ds = (f,v), \quad \forall v \in H_0^1(\mathcal{T}_k), \ t \in I, \tag{5.3.3.1}$$

where B_k was defined by (5.3.2.1), i.e.,

$$\begin{aligned} B_k(u,v) \ := & \sum_{T\in\mathcal{T}_k} (\nabla u, \nabla v)_T \\ & + \sum_{e\in\mathcal{E}_k} \left(\frac{\sigma}{\ell_e}([u],[v])_e - (\langle \partial_n u\rangle,[v])_e - ([u],\langle\partial_n v\rangle)_e \right). \end{aligned}$$

We use the backward Euler scheme to approximate (5.3.3.1), and the trapezoidal rule to discretize the memory term. Then we get the form: Find $w \in H_0^1(\mathcal{T}_k)$, such that

$$A_{\tau,k}(w,v) = (g,v), \quad \forall v \in H_0^1(\mathcal{T}_k), \tag{5.3.3.2}$$

where

$$A_{\tau,k}(w,v) := \tau^{-1}(w,v) + (1 + \frac{1}{2}\tau k_0)B_k(w,v),$$

and

$$(g,v) := (f,v) - (1+\frac{1}{2}\tau k_0)B_k(u^{n-1},v) - \frac{1}{2}\tau k_n B_k(u_0,v) - \sum_{j=1}^{n-1} \tau k_j B_k(u^j,v).$$

The interior penalty discontinuous Galerkin approximation to (5.3.3.2) is formulated as: Find $w_k \in V_k$ such that

$$A_{\tau,k}(w_k,v) = (g,v), \quad \forall v \in V_k. \tag{5.3.3.3}$$

Its error estimate:

$$|||w - w_k|||_{k,\tau} \le Ch_k^{\alpha}(1+\tau^{-1}h_k)^{1/2}||g||_{H^{-1+\alpha}}, \tag{5.3.3.4}$$

can be easily proved by using the techniques in [3].

Define

$$(A_k w_k, v) := A_k(w_k, v) := A_{\tau,k}(w_k, v), \quad \forall w_k,\ v \in V_k. \tag{5.3.3.5}$$

Then (5.3.3.3) can be expressed by

$$A_k w_k = g_k, \tag{5.3.3.6}$$

where

$$g_k \in V_k,\ (g_k, v) := (g, v),\ \forall v \in V_k.$$

From [52], we know that

$$c(1+\frac{1}{2}\tau k_0)(1+\sigma)h_k^{-2} + \tau^{-1} \le \lambda_k \le C(1+\frac{1}{2}\tau k_0)(1+\sigma)h_k^{-2} + \tau^{-1}.$$

Therefore, we derive

$$\begin{aligned} & c\left((1+\frac{1}{2}\tau k_0)(1+\sigma)\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right) \\ \le\ & \lambda_k \le C\left((1+\frac{1}{2}\tau k_0)(1+\sigma)\frac{\lambda_J}{2^{2(J-k)}} + \frac{2^{2(J-k)}-1}{2^{2(J-k)}}\tau^{-1}\right). \end{aligned} \tag{5.3.3.7}$$

In [52] it was also verified that Assumption 5.3.1.2 holds true. Assumption 5.3.1.3 and Lemma 5.3.1.1 are guaranteed, since they only concern the smoother operator R_k. We assume τ satisfies (5.2.2.9). So by recalling (5.3.1.9), we obtain

$$\begin{aligned} & A((I - B_J A_J)w, w) \\ \le\ & C\left(\sum_{k=1}^{J} \frac{\lambda_k^{-\alpha}}{m_k^{\alpha/2}}\right) A(w, w) \\ \le\ & C\left(\frac{1}{1+\sigma+2\sqrt{2(\sigma+1)}}\right)^{\alpha} \frac{h_J^{\alpha}}{m_J^{\alpha/2}} \sum_{k=1}^{J}\left(\frac{2^{\alpha}}{\beta^{\alpha/2}}\right)^{k} A(w, w). \end{aligned}$$

The following theorem is henceforth established.

Theorem 5.3.3.1. *Under Assumption 5.2.1.4, we have the convergence estimate of the CMG Algorithm II for (5.3.3.6)*

$$A((I - B_J A_J)w, w) \leq \begin{cases} C \dfrac{1}{1 - \left(\dfrac{2}{\beta^{1/2}}\right)^{\alpha}} \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \textit{for } \beta > 4, \\ CJ \dfrac{h_J^{\alpha}}{m_J^{\alpha/2}} A(w, w), & \textit{for } \beta = 4, \end{cases}$$

for all $w \in V_J$.

5.4 History of cascading multigrid method

At present there exist three types of multigrid algorithms: the V-cycle algorithm, the W-cycle algorithm and the cascading multigrid algorithm. You may consult Hackbusch [55] and Bramble [19] for background material, and Brandt [21] and Trottenberg *et al.* [115] for references. The cascading algorithm is the new member of the family of multigrid methods. As a distinctive feature, the algorithm does not need nested correction at all and performs more iterations on coarser levels so as to obtain fewer iterations on finer levels. The first publication of this algorithm (Deuflhard [42] in 1994) contained rather convincing numerical results, but no theoretical justification. In 1996, Bornemann and Deuflhard [15] provided a theoretical analysis of this algorithm. In 1998 and 1999, Shi and Xu [104, 105] generalized the idea and applied it to nonconforming finite element method. Later many papers such as ([51], [106], [17], [103], [102], [107], [108], [112], [110], [126], [18], [87]) contributed to this area. In this thesis and [86], we provided an abstract cascasing multigrid, which is more general than [104, 105] and is applicable for the problems with mild regularity in Besov spaces and for the interior penalty discontinuous Galerkin method.

Chapter 6

Future works

In this chapter, we shall mention some future research topics growing out of this thesis.

6.1 Adaptive discontinuous Galerkin time-stepping for (partial) VIDEs with blow-up solutions

6.1.1 VIDEs with blow-up solutions

Consider the semilinear Volterra integro-differential equation:

$$\begin{cases} y'(t) + a(t)y(t) = \mathcal{V}_G(y)(t), & \text{for } t > 0, \\ y(0) = y_0, \end{cases} \tag{6.1.1.1}$$

where $\mathcal{V}_G(y)(t) := \int_0^t k(t-s)G(y(s))ds$. The solution of (6.1.1.1) will blow up in finite time under suitable assumptions on the functions a, k, and G. For example, blow-up will occur in finite time under the following assumptions:

1. $G(y) := y^p, \quad \text{for } p > 1.$
2. For $\tau \geq 0$, k satisfies

$$k \in C^1, \ k(\tau) \geq \text{const.} > 0, \ \text{and } k'(\tau) \leq 0.$$

3. $y_0 \geq 0$ and $0 < -a(t) < \text{const.}$ (for $t > 0$).

(Compare Roberts and Olmstead [94] for the case of Volterra integral equations.

and see Bellout [13] for PVIDEs in Section 6.1.2).

As shown in Stuart and Floater [111], the time-stepping method (one-point collocation method) with fixed temporal mesh is totally inadequate for dealing with ODEs with blow-up solutions. The same is true for the discontinuous Galerkin method on a fixed mesh, and one will resort to techniques for the computation of blow-up problems. There exist various methods for generation of adaptive meshes (e.g., defining the adaptive mesh according to the asymptotic profile of the solution near the blow-up time (if it is known)). Time-stepping based on a posteriori error estimates of DG is definitely one of the efficient approaches for mesh adaptivity.

Another approach for the blow-up problem (6.1.1.1) is the p- or hp-version of the adaptive DG time-stepping method (Brunner and Schötzau [27]). It will likely be very effective in dealing with blow-up equations, especially when the asymptotic behavior of the solution near the blow-up time is known.

6.1.2 PVIDEs with blow-up solutions

Let Ω be a bounded domain in $\mathbb{R}^n$ with piecewise smooth boundary $\partial\Omega$ and

$$Q_t := \Omega \times (0,t), \quad \Gamma_t := \partial\Omega \times (0,t).$$

We consider the PVIDE:

$$u_t = \triangle u + \int_0^t k(t-s)G(u(x,s))ds, \quad \text{in } Q_t, \tag{6.1.2.1}$$

with initial condition

$$u(x,0) = u_0(x) \geq 0, \quad \text{in } \Omega,$$

and boundary condition

$$u(x,t) = 0, \quad \text{on } \Gamma_t.$$

Bellout [13] proved that u blows up in finite time, i.e.,

$$\exists T_b < \infty \text{ such that } \lim_{t \to T_b} \max_{x \in \bar{\Omega}} u(x,t) = \infty,$$

if

$$G \in C^1,\ G(0) > 0;\ G'(\tau) > 0,\ G''(\tau) \geq 0\ (\forall \tau \geq 0), \tag{6.1.2.2}$$

and

$$k \in C^1;\ k(\tau) \geq \text{const.} > 0,\ k'(\tau) \leq 0\ (\forall \tau \geq 0). \tag{6.1.2.3}$$

Problem (6.1.2.1) can be viewed as the generalization of problem (6.1.1.1): because it now also involves a spatial variable, its numerical analysis and computation become much more complicated. Future efficient methods for the computation of (6.1.2.1) are based on *moving meshes* in which the spatial mesh is generated by appropriately chosen moving mesh PDEs (see, e.g., [67, 68, 69, 70], also Bandle and Brunner [7]).

Readers are referred to the book by Samarskii, Galaktionov, Kurdyumov, and Mikhailov [95], an the survey papers by Bandle and Brunner [8] and Souplet [109] for numerous references on theoretical and numerical blow-up.

6.2 The artificial boundary method for PVIDEs on unbounded spatial domains

6.2.1 The artificial boundary method

Consider the following initial-boundary-value problem for the one-dimensional diffusion equation with memory term:

$$u_t + \int_0^t k(x, t-\tau) u(x,\tau) d\tau = \Delta u + f(x,t), \quad x \in \mathbb{R}^1,\ t \in I, \tag{6.2.1.1}$$

$$u|_{t=0} = u_0(x), \quad x \in \mathbb{R}^1, \tag{6.2.1.2}$$

$$u \to 0, \quad \text{as} \quad |x| \to \infty, \tag{6.2.1.3}$$

where $I := [0, T]$, $\Delta u := \partial^2 u / \partial x^2$. Suppose that

(i) the functions f and u_0 are continuous and have compact support:

$$\text{supp}\,\{f(x)\} \subset [0,1], \qquad \text{supp}\,\{u_0(x)\} \subset [0,1];$$

(ii) the kernel k satisfies $k(t,x) = k_0(t)$ when $x \notin (0,1)$. with k_0 continuous or weakly singular.

In order to solve this problem numerically we introduce two artificial boundaries, as follows:

$$\Gamma_1 \quad := \quad \{x = 1 : \; 0 \le t \le T\}, \tag{6.2.1.4}$$

$$\Gamma_0 \quad := \quad \{x = 0 : \; 0 \le t \le T\}. \tag{6.2.1.5}$$

These artificial boundaries divide the given spatial-temporal domain into three sub-domains:

$$\begin{aligned} Q_1 \quad &:= \quad \{(x,t) : \; 1 < x < +\infty, \;\; 0 \le t \le T\}, \\ Q_0 \quad &:= \quad \{(x,t) : \; -\infty < x < 0, \;\; 0 \le t \le T\}, \\ Q_i \quad &:= \quad \{(x,t) : \; 0 \le x \le 1, \;\; 0 \le t \le T\}. \end{aligned}$$

Consider first the restriction of the given initial-boundary-value problem (6.2.1.1)–(6.2.1.3) to the domain Q_1. Because of our assumptions (i) and (ii), $u = u(x,t)$ has to satisfy

$$u_t + \int_0^t k_0(t-\tau) u(x,\tau) d\tau = \Delta u, \quad 1 < x < \infty, \; 0 < t \le T, \tag{6.2.1.6}$$

$$u|_{t=0} = 0, \quad 1 \le x < \infty, \tag{6.2.1.7}$$

$$u \to 0, \quad \text{as} \quad x \to \infty. \tag{6.2.1.8}$$

Using Laplace transform techniques one can show that (see Han, Brunner and Ma [61]) the exact (artificial) boundary conditions on Γ_1 and on Γ_0 are respectively given by

$$\frac{\partial u(1,t)}{\partial x} = -\frac{1}{\sqrt{\pi}} \int_0^t \frac{H(t-\tau)}{\sqrt{t-\tau}} \frac{\partial u(1,\tau)}{\partial \tau}\, d\tau \tag{6.2.1.9}$$

and by

$$\frac{\partial u(0,t)}{\partial x} = +\frac{1}{\sqrt{\pi}} \int_0^t \frac{H(t-\tau)}{\sqrt{t-\tau}} \frac{\partial u(0,\tau)}{\partial \tau}\, d\tau, \tag{6.2.1.10}$$

with $t \in [0, T]$ and with appropriate kernel H.

By the artificial boundary conditions (6.2.1.9) and (6.2.1.10) the original initial-boundary-value problem (6.2.1.1)–(6.2.1.3) can thus be reduced to one defined on the *bounded* spatial-temporal computational domain Q_i:

$$\frac{\partial u}{\partial t} + \int_0^t k(x, t-\tau) u(x,\tau) d\tau = \Delta u + f(x,t), \quad (x,t) \in Q_i, \tag{6.2.1.11}$$

$$u|_{t=0} = u_0(x), \quad 0 \leq x \leq 1, \tag{6.2.1.12}$$

$$\left.\frac{\partial u}{\partial x}\right|_{x=1} = -\frac{1}{\sqrt{\pi}} \int_0^t \frac{H(t-\tau)}{\sqrt{t-\tau}} \frac{\partial u(1,\tau)}{\partial \tau}\, d\tau, \tag{6.2.1.13}$$

$$\left.\frac{\partial u}{\partial x}\right|_{x=0} = +\frac{1}{\sqrt{\pi}} \int_0^t \frac{H(t-\tau)}{\sqrt{t-\tau}} \frac{\partial u(0,\tau)}{\partial \tau}\, d\tau\ . \tag{6.2.1.14}$$

On Q_i the problem (6.2.1.11)–(6.2.1.14) is equivalent to (6.2.1.1)–(6.2.1.3).

Han and Huang [62, 63] proposed an artificial boundary method for the heat equation on unbounded domains. The method focuses on introducing an appropriate computational domain with an artificial boundary and adding the nonlocal boundary condition. Han, Brunner, and Ma [61] used this method to solve linear PVIDEs on unbounded spatial domains in $\mathbb{R}^1$. Work is currently being done on the extension to unbounded domains in $\mathbb{R}^2$ and $\mathbb{R}^3$.

6.2.2 The finite element method and the DG time-stepping for the reduced problems

Since the reduced problem (6.2.1.11)–(6.2.1.14) includes the artificial boundary conditions (6.2.1.9) and (6.2.1.10), it is important to verify the coercivity and continuity of the bilinear form $a(u, v)$ in the error estimate of the finite element approximation to the reduced problem. The reader may consult the following fundamental references:

1. Han and Wu [64]: artificial boundary method for Laplace's equation and linear elastic equations on unbounded domains and error estimates for its finite element approximations.
2. Han and Bao [60]: the finite element approximation of elliptic problems on unbounded domains is formulated on a bounded domain using a nonlocal approximate artificial boundary condition and error estimates are based on the mesh size, the terms used in the approximate artificial boundary condition, and the location of the artificial boundary.
3. Han and Zheng [65]: mixed finite element methods and high-order local artificial boundary conditions for exterior problems of elliptic equations.

The analysis of the finite element method for the reduced problems coming from artificial boundary methods for parabolic PDEs and , especially, PVIDEs, is still at an early stage.

We see from (6.2.1.11)–(6.2.1.14) that the artificial boundary conditions are expressed in time integral form. Hence, it will be necessary to employ suitable quadrature for the boundary conditions while applying the discontinuous Galerkin time-stepping method to the reduced problem. The error analysis of the DG time-stepping for the reduced problem will be based on error estimates of its finite element method in space, the time-stepping error estimates, and the quadrature errors.

6.2.3 Blow-up problems on unbounded spatial domains

An obvious extension of (6.2.1.1)–(6.2.1.3) is the initial-boundary-value problem for the nonlinear diffusion equation with memory term:

$$u_t = \Delta u + \int_0^t k(x, t-\tau) u^p(x, \tau) d\tau, \quad x \in \mathbb{R}^n,\ t > 0,\ p > 1, \quad (6.2.3.1)$$

$$u|_{t=0} = u_0(x), \quad x \in \mathbb{R}^n, \quad (6.2.3.2)$$

$$u \longrightarrow 0, \quad \text{as} \quad |x| \longrightarrow \infty. \quad (6.2.3.3)$$

The blow-up property for (6.2.3.1)–(6.2.3.3) with unbounded spatial domains is not yet known, in contrast to that for parabolic PDEs,

$$u_t = \triangle u + u^p,$$

compare Fujita [48] (in $\mathbb{R}^n$), Bandle and Levine [9] (for sectorial domains), Bandle and Brunner [8], and their references. For the purpose of computation of the problem (6.2.3.1)–(6.2.3.3), we shall introduce the artificial boundaries and the corresponding artificial boundary conditions. We have be very careful on the choice of the location of the artificial boundary, otherwise it is possible that some blow-up points are not included into the computational domains. Is it true that the blow-up property of the blow-up of the reduced problem with nonlinear artificial boundary conditions is the same as that of the original problem (6.2.3.1)–(6.2.3.3)? If we have not established these corresponding analysis, how can we numerically detect the blow-up and determine the suitable location of the artificial boundaries?

Bibliography

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, John Wiley & Sons, New York, 2000.

[3] D.N. Arnold, An interior penalty finite element method with discontinuous elements, *SIAM J. Numer. Anal.*, 19 (1982), 742–760.

[4] O. Axelsson and W. Layton, A two-level discretization of nonlinear boundary value problems, *SIAM J. Numer. Anal.*, 33 (1996), 2359–2374.

[5] I. Babuška and M. Suri, the p and hp-versions of the finite element method: basic principles and properties, *SIAM Review*, 36 (1994), 578–632.

[6] C. Bacuta, J.H. Bramble, and J. Xu, Regularity estimates for elliptic boundary value problems in Besov spaces, *Math. Comp.*, 72 (2002), 1577–1595.

[7] C. Bandle and H. Brunner, Numerical analysis of semilinear parabolic problems with blow-up solutions, *Rev. Real Acad. Cienc. Exact. Fís. Natur. Madrid*, 88 (1994), 203–222.

[8] C. Bandle and H. Brunner, Blowup in diffusion equations: a survey, *J. Comput. Appl. Math.*, 97 (1998), 3–22.

[9] C. Bandle and H.A. Levine, Fujita type phenomena for reaction-diffusion equations with convection like terms, *Differential Integral Equations*, 7 (1994), 1169–1193.

[10] R.E. Bank and T. Dupont, An optimal order process for solving finite element equations, *Math. Comp.*, 36 (1980), 35–51.

[11] D.M. Bedivan, A two-grid method for solving elliptic problems with inhomogeneous boundary conditins, *Comput. Math. Appl.*, 29 (1995), 59–66.

[12] A. Bellen and M. Zennaro, *Numerical Methods for Delay Differential Equations*, Oxford University Press, New York, 2003.

[13] H. Bellout, Blow-up solutions of parabolic equations with nonlinear memory, *J. Differential Equations*, 70 (1987), 42–68.

[14] T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, and P. Krysl, Meshless methods: an overview and recent developments, *Comput. Methods Appl. Mech. Engrg.*, 139 (1996), 3–47.

[15] F. Bornemann and P. Deuflhard, The cascadic multigrid method for elliptic problems, *Numer. Math.*, 75 (1996), 135–152.

[16] F.A. Bornemann and P. Deuflhard, Cascadic multigrid methods, in *Domain Decomposition Methods in Sciences and Engineering (Beijing, 1995)*, 205–212, Wiley, Chichester, 1997.

[17] D. Braess and W.A. Dahmen, Cascade multigrid algorithm for the Stokes equations, *Numer. Math.*, 82 (1999), 179–192.

[18] D. Braess, P. Deuflhard, and K. Lipnikov, A subspace cascadic multigrid method for mortar elements, *Computing*, 69 (2002), 205–225.

[19] J.H. Bramble, *Multigrid Methods*, John Wiley & Sons, New York, 1993.

[20] A. Brandt and C.W. Cryer, Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems, *SIAM J. Sci. Stat. Computing*, 4 (1983), 655–684.

[21] A. Brandt, Multiscale scientific computation: review 2001, in *Multiscale and Multiresolution Methods*, 3–95, Lect. Notes Comput. Sci. Eng., 20, Springer, Berlin, 2002.

[22] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 2002.

[23] H. Brunner, *Collocation Methods for Volterra Integral and Related Functional Differential Equations*, Cambridge University Press, Cambridge, 2004 (to appear).

[24] H. Brunner, The numerical solution of weakly singular Volterra integral equations by collocation on graded meshes, *Math. Comp.*, 45 (1985), 417–437.

[25] H. Brunner, Q.Y. Hu, and Q. Lin, Geometric meshes in collocation methods for Volterra integral equations with proportional delays, *IMA J. Numer. Anal.*, 21 (2001), 783–798.

[26] H. Brunner and P.J. van der Houwen, *The Numerical Solution of Volterra Equations*, North-Holland, Amsterdam, 1986.

[27] H. Brunner and D. Schötzau, An *hp*-error analysis of the DG method for parabolic Volterra integro-differential equations with weakly singular kernels, (to appear).

[28] H. Brunner and W.K. Zhang, Primary discontinuities in solutions for delay integro-differential equations, *Methods Appl. Anal.*, 6 (1999), 525–533.

[29] H. Brunner and J.T. Ma, Discontinuous Galerkin approximation for Volterra integro-differential equations with variable delays and weakly singular kernels, (to appear).

[30] J.C. Butcher, *Numerical Methods for Ordinary Differential Equations (Second Edition)*, Wiley, New York, 2003.

[31] B. Cockburn, G.E. Karniadakis, and C.-W. Shu, *Discontinuous Galerkin Methods*, Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin - Heidelberg, 2000.

[32] J.M. Chadam, A. Peirce, and H.M. Yin, The blowup property of solutions to some diffusion equations with localized nonlinear reactions, *J. Math. Anal. Appl.*, 169 (1992), 313–328.

[33] C. Corduneanu, *Integral Equations and Applications*, Cambridge University Press, Cambridge, 1991.

[34] C. Chen and T. Shih, *Finite Element Methods for Integro-Differential Equations*, Series on Applied Math., Vol. 9, World Scientific, Singapore, 1998.

[35] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[36] M. Crouzeix, V. Thomée, and L. Wahlbin, Error estimates for spatially discrete approximations of smilinear parabolic equations with initial data of low regularity. *Math. Comp.*, 53 (1989), 25-41.

[37] J.M. Cushing, *Integro-Differential Equations and Delay Models in Population Dynamics*, Lecture Notes Biomath., Vol. 20, Springer-Verlag, New York, 1977.

[38] C.N. Dawson and M.F. Wheeler, Two-grid methods for mixed finite element approximations of nonlinear parabolic equations, *Contemp. Math.*, 180 (1994), 191–203.

[39] C.N. Dawson, M.F. Wheeler, and C.S. Woodward, A two-grid finite difference scheme for nonlinear parabolic equations, *SIAM J. Numer. Anal.*, 35 (1998), 435–452.

[40] M. Delfour and F. Dubeau, Discontinuous polynomial approximations in the theory of one-step, hybrid and multistep methods for nonlinear ordinary differential equations, *Math. Comp.*, 47 (1986), 169–189.

[41] M. Delfour, W. Hager, and F. Trochu, Discontinuous Galerkin methods for ordinary differential equations, *Math. Comp.*, 36 (1981), 455–473.

[42] P. Deuflhard, Cascadic conjugate gradient methods for elliptic partial differential equations: algorithm and numerical results, in *Domain Decomposition Methods in Scientific and Engineering Computing* (University Park, PA, 1993), 29–42, Contemp. Math., 180, Amer. Math. Soc., Providence, RI, 1994.

[43] D. Estep, A posteriori error bounds and global error control for approximation of ordinary differential equations, *SIAM J. Numer. Anal.*, 32 (1995), 1–48.

[44] H. Engels, *Numerical Quadrature and Cubature*, Academic Press, 1980.

[45] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, 1996.

[46] G. Fairweather, Spline collocation methods for a class of hyperbolic partial integro-differential equations, *SIAM J. Numer. Anal.*, 31 (1994), 444–460.

[47] J. Fröhlich and J. Lang, Two-dimensional cascadic finite element computations of combustion problems, *Comput. Methods Appl. Mech. Engrg.*, 158 (1998), 255–267.

[48] H. Fujita, On the blowing up of solutions of the Cauchy problem for $u_t = \triangle u + u^{1+\alpha}$, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 13 (1966), 109–124.

[49] Y. Fujita, Integral equation wich interpolates the heat equation and the wave equation, *Osaka J. Math.*, 27 (1990), 309–321; 28 (1990), 797–804.

[50] C.W. Gear, Hybrid methods for initial value problems in ordinary differential equations, *SIAM J. Numer. Anal.*, 2 (1965), 69–86.

[51] L.V. Gilyova and V.V. Shaidurov, A cascadic multigrid algorithm in the finite element method for the plane elasticity problem, *East-West J. Numer. Math.*, 5 (1997), 23–34.

[52] J. Gopalakrishnan and G. Kanschat, A multilevel discontinuous Galerkin method, *Numer. Math*, 95 (2003), 527–550.

[53] G. Gripenberg, S.-O. Londen, and O. Staffans, *Volterra Integral and Functional Equations*, Cambridge University Press, Cambridge, 1990.

[54] W. Gui and I. Babuška, The h, p and h-p versions of the finite element method in one dimension, Part 1: the error analysis of the p-version; Part 2: the error analysis of the h and h-p version; Part 3: the adaptive h-p version, *Numer. Math.*, 40 (1986), 577–612; 613–657; 659–683.

[55] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, New York - Berlin - Heidelberg, 1985.

[56] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer-Verlag, New York, 1994.

[57] E. Hairer, S.P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations*, Springer-Verlag, Berlin, 1993.

[58] J.K. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1977.

[59] J.K. Hale and S.M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, New York, 1993.

[60] H.D. Han and W.Z. Bao, Error estimates for the finite element approximation of problems in unbounded domains, *SIAM J. Numer. Anal.*, 37 (2000), 1101–1119.

[61] H.D. Han, H. Brunner, and J.T. Ma, The numerical solution of parabolic Volterra integro-differential equations on unbounded spatial domains, (to appear).

[62] H.D. Han and Z.Y. Huang, A class of artificial boundary conditions for heat equation in unbounded domains, *Comput. Math. Applic.*, 43 (2002), 889–900.

[63] H.D. Han and Z.Y. Huang, Exact and approximation boundary conditions for the parabolic problems on unbounded domains, *Comput. Math. Applic.*, 44 (2002), 655–666.

[64] H.D. Han and X.N. Wu, Approximation of infinite boundary condition and its application to finite element methods, *J. Compt. Math.*, 3 (1985), 181–192.

[65] H.D. Han and Ch. Zheng, Mixed finite element method and high-order local artificial boundary conditions for exterior problems of elliptic equation, *Comput. Methods Appl. Mech. Engrg.*, 191 (2002), 2011–2027.

[66] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, New York, 1987.

[67] W. Huang, Y. Ren, and R.D. Russell, Moving mesh partial differential equations (MMPDES) based on the equidistribution principle, *SIAM J. Numer. Anal.*, 31 (1994), 709–730.

[68] W. Huang and R.D. Russell, A moving collocation method for solving time dependent partial differential equations, *Appl. Numer. Math.*, 20 (1996), 101–116.

[69] W. Huang and R.D. Russell, Analysis of moving mesh partial differential equations with spatial smoothing, *SIAM J. Numer. Anal.*, 34 (1997), 1106–1126.

[70] W. Huang and D.M. Sloan, A simple adaptive grid method in two dimensions, *SIAM J. Sci. Comput.*, 15 (1994), 776–797.

[71] Y. Huang, Z. Shi, T. Tang, and W. Xue, A multilevel successive iteration method for nonlinear elliptic problems, *Math. Comp.*, 73 (2004), 525-539.

[72] Y. Huang and W.M. Xue, Convergence of finite element approximations and multilevel linearization for Ginzburg-Landau model of d-wave superconductors, *Adv. Comput. Math.*, 17 (2002), 309–330.

[73] B.L. Hulme, Discrete Galerkin and related one-step methods for ordinary differential equations, *Math. Comp.*, 26 (1972), 881–891.

[74] B.L. Hulme, One-step piecewise polynomial Galerkin methods for initial value problems, *Math. Comp.*, 26 (1972), 415–426.

[75] M. Iida, Exponentially asmptotic stability for a certain class of semilinear Volterra diffusion equations, *Osaka Math. J.*, 28 (1991), 411–440.

[76] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, New York, 1987.

[77] C. Johnson, Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations, *SIAM J. Numer. Anal.*, 25 (1988), 908–926.

[78] C. Johnson, S. Larsson, V. Thomée, and L. Wahlbin, Error estimates for spatially discrete approximations of smilinear parabolic equations with nonsmooth initial data, *Math. Comp.*, 49 (1987), 331-357.

[79] A. Kufner, O. John, and S. Fŭcik, *Function Spaces*, Monographs and Textbooks on Mechanics of Solids and Fluids Mechanics: Analysis, Noordhoff International Publishing, Prague. 1977.

[80] S. Larsson, V. Thomée, and L.B. Wahlbin, Numerical solution of parabolic integro-differential equations by the discontinuous Galerkin method, *Math. Comp.*, 67 (1998), 45–71.

[81] W. Layton and W. Lenferink, Two-level Piard and modified Picard methods for the Navier-Stokes equations, *Appl. Math. Comp.*, 69 (1995), 263–274.

[82] P. Lesaint and P.A. Raviart, On a finite element method for solving the neutron transport equation, in *Mathematical Aspects of Finite Elements in Partial Differential Equations (C. de Boor, 1974)*, 89–145, Academic Press, 1974.

[83] J.T. Ma and H. Brunner, A posteriori error estimates of discontinuous Galerkin methods for nonstandard Volterra integro-differential equations, *IMA J. Numer. Anal.*, (to appear).

[84] J.T. Ma and H. Brunner, Cascading multilevel discretization algorithms for semilinear parabolic problems, (to appear).

[85] G.R. Liu, *Mesh Free Methods: Moving beyond the Finite Element Method*, CRC Press, New York, 2003.

[86] J.T. Ma and H. Brunner, Abstract cascading multigrid preconditioners in Besov spaces, (to appear).

[87] J.T. Ma, Z.C. Shi, and J.P. Zeng, Multilevel successive iteration method for variational inequality problems, *J. Hunan Univ.*, 30 (2003), 5–7.

[88] M. Marion and J.C. Xu, Error estimates on a new nonlinear Galerkin method based on two-grid finite elements, *SIAM J. Numer. Anal.*, 32 (1995), 1170–1184.

[89] R.K. Miller, On Volterra's polulation equation, *SIAM J. Appl. Math.*, 14 (1966), 446–452.

[90] R.K. Miller, *Nonlinear Volterra Integral Equations*, Benjamin, Menlo Park, Calif., 1971.

[91] A.K. Pani and R.K. Sinha, Error estimates for semidiscrete Galerkin approximation to time dependent parabolic integro-differential equation with nonsmooth data, *Calcolo*, 37 (2000), 181–205.

[92] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer-Verlag, 1994.

[93] J.R. Rice, On the degree of convergence of nonlinear spline approximation, *Approximation with Special Emphasis on Spline Functions* (I.J. Schoenberg, Ed.), Academic Press, New York, 1969.

[94] C.A. Roberts and W.E. Olmstead, Growth rates for blow-up solutions of nonlinear Volterra equations, *Quart. Appl. Math.*, 54 (1996), 153–159.

[95] A.A. Samarskii, V.A. Galaktionov, S.P. Kurdyumov, and A.P. Mikhailov, *Blow-up in Quasilinear Parabolic Equations*, New York, 1995.

[96] D. Schötzau and C. Schwab, An *hp* a priori error analysis of the DG time-stepping method for initial value problems, *Calcolo*, 37 (2000), 207–232.

[97] D. Schötzau and C. Schwab, Time discretization of parabolic problem by the *hp*-version of the discontinuous Galerkin finite element method, *SIAM J. Numer. Anal.*, 38 (2000), 837–875.

[98] C. Schwab, *p- and hp-Finite Element Methods*, Oxford University Press New York, 1998.

[99] S. Shaw and J.R. Whiteman, Discontinuous Galerkin method with a posteriori $L_p(0, t_i)$ error estimate for second-kind Volterra problems, *Numer. Math.*, 74 (1996), 361–383.

[100] S. Shaw and J.R. Whiteman, Adaptive space-time finite element solution for Volterra equations arising in viscoelasticity problems, *J. Comp. Appl. Math.*, 125 (2000), 337–345.

[101] V. Shaidurov, Some estimates of the rate of convergence for the cascadic conjugate gradient method, *Comp. Appl. Math.*, 31 (1996), 161–171.

[102] V. Shaidurov and G. Timmermann, A cascadic multigrid algorithm for semilinear indefinite elliptic problems, in International GAMM-Workshop on Multigrid Methods (Bonn, 1998), *Computing*, 64 (2000), 349–366.

[103] V. Shaidurov and L. Tobiska, The convergence of the cascadic conjugate-gradient method applied to elliptic problems in domains with re-entrant corners. *Math. Comp.*, 69 (2000), 501–520.

[104] Z. Shi and X. Xu, Cascadic multigrid for elliptic problems, *East-West J. Numer. Math.*, 6 (1998), 309–318.

[105] Z. Shi and X. Xu, Cascadic multigrid for elliptic problems, *East-West J. Numer. Math.*, 7 (1999), 199–211.0

[106] Z. Shi and X. Xu, Cascadic multigrid for the plate bending problem, *East-West J. Numer. Math.*, 6 (1998), 137–153.

[107] Z. Shi and X. Xu, Cascadic multigrid for parabolic problems, *J. Comp. Math.*, 18 (2000), 450–459.

[108] Z. Shi and X. Xu, A new cascadic multigrid, *Sci. China Ser. A*, 44 (2001), 21–30.

[109] Ph. Souplet, Blow-up in nonlocal reaction-diffusion equations, *SIAM J. Math. Anal.*, 29 (1998), 1301–1334.

[110] R. Stevenson, Nonconforming finite elements and the cascadic multi-grid method, *Numer. Math.*, 91 (2002), 351–387.

[111] A.M. Stuart and M.S. Floater, On the computation of blow-up, *Europ. J. Appl. Math.*, 1 (1990), 47–71.

[112] G. Timmermann, A cascadic multigrid algorithm for semilinear elliptic problems, *Numer. Math.*, 86 (2000), 717-731.

[113] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, New York, 1997.

[114] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, North-Holland, Amsterdam, 1978.

[115] U. Trottenberg, C.W. Oosterlee, and A. Schüller, *Multigrid*, with contributions by A. Brandt, P. Oswald, and K. Stüben, Academic Press, Inc., San Diego, CA, 2001.

[116] T. Utnes, Two-grid finite element formulations of the incompressible Navier-Stokes equations, *Comm. Numer. Methods Engrg.*, 34 (1997), 675–684.

[117] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*, Dover Publication, Inc., New York, 1959.

[118] V. Volterra, Sulle equazioni integro-differenziali, *Rend. Accad. Lincei* (5). 18 (1909), 167–174.

[119] J.C. Xu, Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.*, 33 (1996), 1759-1777.

[120] J.C. Xu, A novel two-grid method for semilinear elliptic equations, *SIAM J. Sci. Comp.*, 15 (1994), 231-237.

[121] J.C. Xu, A new class of iterative methods for nonselfadjoint or indefinite problems, *SIAM J. Numer. Anal.*, 29 (1992), 303–319.

[122] J.C. Xu, *Theory of Multilevel Methods*, Rep AM-48, Dept. Math., Penn State, 1989.

[123] J.C. Xu and A.H. Zhou, Local and parallel finite element algorithms based on two-grid discretizations, *Math. Comp.*, 69 (1999), 881–909.

[124] J.C. Xu and A.H. Zhou, Local and parallel finite element algorithms based on two-grid discretizations for nonlinear problems, *Adv. Comput. Math.*, 14 (2001), 293–327.

[125] M. Zennaro, Delay differential equations: theory and numerics, in *Theory and Numerics of Ordinary and Partial Differential Equations*, M. Ainsworth, et al., ed., Clarendon Press, Oxford, 1995, 291–333.

[126] J.P. Zeng and J.T. Ma, A cascadic multigrid method for a kind of one-dimensional elliptic variational inequality, *J. Hunan Univ.*, 28 (2001), 1–5.